Four Essays in Asset Pricing

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Ole Wilms
from Germany

approved in April 2016 at the request of
Prof. Dr. Karl Schmedders
Prof. Dr. Felix Kübler

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, April 06, 2016

Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

I am heavily indebted and very grateful to my supervisors Felix Kübler and Karl Schmedders, as well as to Ken Judd, Thomas Lontzek and Walter Pohl for their teaching, guidance and support on this project. Also I would like to give special thanks to Paula and my parents for their continuous support and to Gregor Reich for the helpful discussions on the subject.

Contents

Part I

Introduction

General Introduction

Asset pricing models have advanced a great deal in the past decades. Early work in the 1980s has shown that in a standard representative agent economy with time separable preferences an unrealistic degree of risk aversion is needed to explain large-scale features of financial markets, such as the high equity premium (Grossman and Shiller (1981), Hansen and Singleton (1982) and Mehra and Prescott (1985)). In the subsequent years there have been many attempts to modify and extend the standard model to be consistent with asset pricing data.[1] While the success was rather mixed, in the early 2000s there emerged a new strand of literature, the so-called long-run risk models (Bansal and Yaron (2004)). By combining highly persistent shocks to mean consumption growth with recursive preferences that are calibrated to model a preference for the early resolution of risks, the authors had great success in explaining long-standing asset pricing puzzles. Thereafter many researchers have followed their approach and extended the framework making it the workhorse model in the recent asset pricing literature (e.g. Hansen, Heaton, and Li (2008), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Bansal and Shaliastovich (2013) or Drechsler and Yaron (2011) among others).

The major goal of this thesis is to contribute to this literature by analyzing several topics in consumption-based asset pricing.[2] In Chapter 1 we analyze long-run risk models with regard to the existence of solutions as well as the influence of higher-order dynamics. In Chapter 2 we show how to include agent heterogeneity in the model to understand the impact of differences among agents on model outcomes. Chapter 3 analyzes the different influences of temporary and permanent shocks on financial markets and in Chapter 4 we compare methods to solve asset pricing models.

We address these topics in the following way:

Chapter 1 examines the existence of solutions and the effects of higher-order dynamics on equilibrium outcomes in long-run risk models. Long-run risk models combine complex specifications for the exogenous driving forces of the economy with recursive preferences of Epstein-Zin type (Epstein and Zin (1989) and Weil (1990)) that are potentially highly non-linear. When solving

[1]See for example the habit literature by Abel (1990), Constantinides (1990) and Campbell and Cochrane (1999) or the literature on rare disasters by Rietz (1988) and Barro (2006).

[2]There are two further projects that I worked on during the time of the dissertation. In Reich and Wilms (2015) we show how to use flexible grids for the estimation of dynamic models and in Lontzek, Narita, and Wilms (2015) we examine optimal resource extraction policies under tipping point risks. As both topics are fundamentally different from the asset pricing chapters in this thesis, they are not included in the dissertation.

such models two questions naturally arise: First, does a solution to the model exist and second, if yes, how do we reliably compute it?

To answer the first question, we use results on the existence of solutions under the restrictive assumption of constant relative risk aversion (CRRA) preferences, where the investor is neutral about the resolution of risks. Under this assumption, de Groot (2015) provides closed-form solutions for the Bansal and Yaron (2004) model as well as a formal existence theorem. We prove existence for the general case of Epstein-Zin preferences using a simple relative result. If the model has a solution under the special assumption of CRRA preferences, then it also has a solution if the agent has a preference for the early resolution of risks for the same specification. Hence, as de Groot (2015) proves existence for the CRRA version of the long-run risk model, existence with Epstein-Zin utility and a preference for the early resolution of risks follows.

So we know that a solution exists, but how do we compute it? The most commonly used approach to solve long-run risk models is log-linearization. Based on the Campbell and Shiller (1988a) present-value relation, Bansal and Yaron (2004) introduce a simple linearizing method that allows to compute approximate closed-form solutions of the model. This approach is particularly attractive as it allows to draw clear-cut conclusions about the mechanisms generating the asset pricing dynamics. While this is desirable from an explanatory point of view, linearization will, by construction, miss the influence of higher-order effects. In the chapter we show that these effects can be large and significant. For this we examine the influence of higher-order dynamics in six recent long-run risk models. We report that the non-linearities strongly dependent on the persistence of the state processes. As the main feature of long-run risk models are highly persistent shocks to mean consumption growth, using linearized solutions potentially introduces large errors. In particular we find that approximation errors are as large as 50% for some key model moments, and log-linearization can even introduce errors in qualitative conclusions predicting a wrong sign of the slope of the yield curve. We conclude that more sophisticated methods should be used to solve long-run risk models that feature highly persistent state processes. One potential candidate are projection methods as we show in the fourth chapter of this thesis.

Chapter 2 presents a general framework to analyze the influence of agent heterogeneity in asset pricing models with recursive preferences. Under the standard assumption of CRRA utility, the influence of agent heterogeneity on model outcomes is well understood. For example, Sandroni (2000) and Blume and Easley (2006) analyze the influence of differences in beliefs on market selection and show that agents with systematically wrong beliefs about future quantities will eventually be driven out of the market. Likewise, Judd, Kubler, and Schmedders (2003) find that trading volume in a simple asset pricing model with CRRA preferences is zero and hence there must be others reasons than differences in the utility parameters to explain the large trading volume observed in the financial data.

However, recent work by Borovička (2015) and Branger, Dumitrescu, Ivanova, and Schlag (2011) suggest that these findings do not hold true under the general assumption of recursive preferences. As recursive preferences are a key feature to explain asset pricing data in long-run risk models, it is of great importance to obtain a better understanding of how agent heterogeneity affects market outcomes in this model framework.

In the chapter we show how such asset pricing models with heterogeneous agents and recursive utility can be solved. The main contribution of the chapter is to derive first-order conditions and present a computational approach, based on projection methods, to compute the equilibrium functions. We apply the method to solve the Bansal and Yaron (2004) model with two agents. A full qualitative and quantitative analysis of the model is beyond the scope of the chapter, but the example serves to demonstrate the solution approach. In addition, the results give the reader insigtphts about the potential influences of agent heterogeneity on model outcomes and motivate further research. Research questions that can be addressed using the methodology presented in the chapter are for example: What are the implications of agent heterogeneity on long-run survival, the wealth distribution, the pricing kernel or aggregate financial market outcomes?

In Chapter 3 we analyze the different implications of temporary and permanent shocks to consumption for asset pricing quantities. The standard assumption in modern asset pricing models (e.g. Bansal and Yaron (2004) or Hansen, Heaton, and Li (2008)) is that shocks to consumption are permanent. We show that by including temporary shocks in a canonical asset pricing model, many asset pricing features can be explained with standard time separable preferences of CRRA type and a reasonable degree of risk aversion.

Given the limited amount of time-series data, it is difficult to distinguish shocks to consumption that are permanent from those that are temporary but very persistent, or even a mix of both kinds of shocks. Fortunately, our findings do not rely on the absence of permanent shocks, but the results hold true as long as temporary shocks represent a significant fraction of the aggregate risk in the economy. In particular we find that temporary shocks to consumption can generate a high risk premium together with a low risk-free rate, volatile stock prices and predictability of returns with a risk aversion coefficient below ten. We also run several tests on the model to verify its consistency with the financial data. For example we report that, in line with the data, most of the variation in the price-dividend ratio is driven by changes in expected returns rather than changes in expected dividends and that the model meets the Hansen-Jagannathan bounds.

In the fourth and last chapter we compare different solution methods for solving asset pricing models with Epstein-Zin preferences. As Chapters 1-3 show, closed-form solutions are rarely available for modern asset pricing models that have become more and more complex. So what should the researcher do if closed-form solution for the model of interest cannot be obtained? In this chapter we describe a projection based approach to solve asset pricing models with Epstein-Zin preferences and compare it to alternative computational procedures.

The commonly used approach in the class of asset pricing models we consider is log-linerization. The advantages of the method are obvious. It is easy to implement, very fast, and in many cases even allows for approximate closed-form solutions (Eraker (2008), Eraker and Shaliastovich (2008)). However, in line with the results from Chapter 1, we find large errors in key quantities for the recent calibration of the long-run risk model by Bansal, Kiku, and Yaron (2012a), and the accuracy of the method strongly depends on the model specification.

The second class of methods we consider are discretization methods (Tauchen (1986), Tauchen and Hussey (1991)) that have for example been used in work by Guvenen (2009), Heaton and Lucas (1996), Heaton and Lucas (2000) or Hansen, Heaton, and Yaron (1996). Discretization methods may, at least in theory, guarantee convergence to the true solution but their accuracy depends highly on the model parameters and they show severe difficulties for highly persistent state processes. This implies that a large number of discretization nodes is needed to obtain a reasonable accuracy. Since computation time increases dramatically with the number of nodes, particularly in higher dimensions, the discretization methods are all but practical for modern asset pricing models.

As we show in the chapter, the implementation of projection methods to solve asset pricing models is somewhat more challenging than the other approaches. But once invested, the effort comes with a high reward. The method proves to be highly accurate and the performance is robust with regard to changes in the model parameters. Even for a low degree approximation it provides accurate solutions for the two-dimensional long-run risk model of Bansal and Yaron (2004), while it is only slightly slower compared to the linearization, suggesting that it can also be used to solve higher dimensional models. We conclude that commonly used methods like linearization and discretization potentially introduce large approximation errors while projection methods provide very accurate and robust solutions. Hence, future research should rather use more sophisticated methods, like the projection approach, to reduce errors that can have significant influence on model outcomes.

Part II

Four Essays in Asset Pricing

Essay 1

Existence and Non-Linearities in Asset Pricing

Higher-Order Effects in Asset Pricing Models with Long-Run Risks[1]

Karl Schmedders
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
karl.schmedders@business.uzh.ch

Walter Pohl
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
walter.pohl@business.uzh.ch

Ole Wilms
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
ole.wilms@business.uzh.ch

November 2015

Abstract

This paper analyzes both the existence of solutions to long-run risk asset pricing models, as well as the practicality of approximating this solutions by the Campbell-Shiller log-linearization. We prove a simple relative existence result that is sufficient to show that the original Bansal-Yaron model has a solution. Log-linearization fares less well: we find that for very persistent processes the approximation errors in model moments can be as large as 50%, and can get such basic facts wrong as the direction of the yield curve.

Keywords: Asset pricing; Long-run risk; Log-linearization; Nonlinear dynamics.

JEL Classification: G11, G12.

Note: A version of this paper has been submitted to Econometrica.

[1]We are indebted to Lars Hansen, Ken Judd, and Martin Lettau for helpful discussions on the subject. We thank seminar audiences at the University of Zurich, the Becker Friedman Institute at the University of Chicago and at various conferences for comments. Karl Schmedders gratefully acknowledges financial support from the Swiss Finance Institute. Ole Wilms gratefully acknowledges financial support from the Zürcher Universitätsverein.

1.1 Introduction

This paper presents an analysis of the existence and the computation of solutions to asset-pricing models that feature long-run risk. We provide a formal existence theorem for the basic Bansal and Yaron (2004) long-run risk model. We then show numerically that the model solutions are potentially very nonlinear, and that for many plausible choices of parameters and exogenous processes the errors introduced by linearization are economically significant. In fact, for very persistent processes the approximation errors in model moments can be as large as 50%, and can get such basic facts wrong as the direction of the yield curve. The increasing complexity of state-of-the-art asset-pricing models leads to complex nonlinear equilibrium functions with considerable curvature which in turn can have sizable economic implications. Therefore, these models require numerical solution methods, such as the projection methods employed in this paper, that can adequately describe the higher-order equilibrium features.

Asset-pricing models have become increasingly complex over the last three decades. The first generation of such models, developed in the 1980s (Grossman and Shiller (1981), Hansen and Singleton (1982), Mehra and Prescott (1985)), proved inadequate in explaining large-scale features of financial markets, such as the high equity premium and the low risk-free rate. As the literature on asset-pricing evolved and matured over time, researchers added more and more complex features to their models with, among others, incomplete markets in form of uninsurable income risks, frictions such as borrowing or collateral constraints, time-varying risk aversion, and heterogeneous expectations. While these additional features had varying degrees of success, recently the new generation of long-run risks models (e.g. Bansal and Yaron (2004) or Hansen, Heaton, and Li (2008)) with their interplay of long-run risks, stochastic volatility, and recursive preferences have had considerably more success in resolving long-standing asset pricing puzzles.

An important part of the appeal of the long-run risk model is that Bansal and Yaron (2004) introduce a simple linearized solution method based on the Campbell and Shiller (1988a) present-value relation. Long-run risk models feature both highly nonlinear preference structures as well as complex specifications for the exogenous driving forces of the economy. To handle the complexity, researchers must resort to some sort of numerical approximation procedure to make their models tractable. Bansal and Yaron showed that for their original model the log price-dividend ratio could be well-approximated by a linear function of the underlying shocks. The linearized Campbell-Shiller solution, which adjusts for the impact of risk on the average price-dividend ratio, is a considerable advance over the traditional method of log-linearizing around the deterministic steady state, which is known to provide a poor approximation for Epstein-Zin preferences (e.g. Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012), Juillard (2011) or de Groot (2013)).

But time marches on, and researchers have moved with it. By its very nature, a log-linear approximation will miss higher-order effects. Can we always safely ignore these higher-order effects? To answer this question, we examine higher-order dynamics in five additional recent studies, the newly calibrated version of the Bansal and Yaron (2004) model by Bansal, Kiku, and Yaron (2012a), the extensive calibration study of Schorfheide, Song, and Yaron (2014), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015) and the work on real and nominal bonds of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013).

We show that the errors introduced by the Campbell-Shiller approximation can be large and economically significant. For example, Bansal, Kiku, and Yaron (2012a) recalibrate the original Bansal and Yaron (2004) model to have more persistent shocks to stochastic volatility. We find that for this calibration the log-linearization introduces approximation errors as large as 22% for key quantities such as the equity premium or the volatility of price-dividend ratio. Schorfheide, Song, and Yaron (2014) perform a Bayesian estimation of the model using the same approximation, and find evidence for a higher persistence for long-run risk. In this case we find approximation errors as large as 50% for some key model moments. In general, highly persistent processes lead to solutions that are highly nonlinear, and thus economically relevant approximation errors. Log-linearization can even introduce errors in qualitative conclusions. For example, under high persistence log-linearization can actually invert the slope of the yield curve in the nominal bond models of Bansal and Shaliastovich (2013) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010).

As an alternative solution procedure, we use the projection method to solve the nonlinear fixed-point equation for the wealth portfolio. It is known (Atkinson (1992)) that if the fixed-point equation has a solution, then under weak conditions the projection method will converge to a solution. This leads us to consider the question of the existence of a solution. Marinacci and Montrucchio (2010) and Hansen and Scheinkman (2012) prove general theorems about the existence of solutions, but for the types of models considered by the long-run risk literature, the existence of a solution is quite delicate, and depends on specific values of both the preference and exogenous process parameters. We prove a simple *relative* result – if the model has a solution under CRRA preferences for a particular exogenous process specification, then it has a solution for an investor that prefers early resolution for the same specification. For investors that prefer late resolution, the implication goes the other way. We then adapt the results in de Groot (2015) to show that the CRRA version of the long-run risk model has a solution, from which existence for Bansal and Yaron (2004) follows.

Summarizing, by construction, log-linearizing the model as it is commonly done in the asset pricing literature misses higher-order dynamics by construction. If the driving factors of the economy are of low persistence or the risk aversion of the representative agent is low, these dynamics will have a negligible influence on equilibrium outcomes. However, the combination of highly persistent processes, together with recursive preferences and a risk aversion significantly larger than one, can introduce strong non-linear dynamics to the model. We show that these errors have a large impact on key financial statistics in many recent asset pricing studies and introduce a bias to the model parameters when it comes to estimation or calibration of the model. Therefore, in the future more sophisticated solution methods should be used, as for example projection methods, that can account for higher-order dynamics.

The paper is organized as follows. Section 1.2 describes the general model framework that is used throughout the paper. In Section 1.3 we provide a formal theorem for the existence of solution in the economy and analyze the key factors determining existence. Afterwards we examine the effect of higher-order dynamics in six recent asset pricing studies in Section 1.4. Section 1.5 concludes.

1.2 Model Framework

We consider a standard asset pricing model with a representative agent and recursive preferences as in Epstein and Zin (1989) and Weil (1990). Indirect utility at time t, V_t, is given recursively as

$$V_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} . \tag{1.1}$$

In this parametrization, C_t is consumption, δ is the time discount factor, γ determines the level of relative risk aversion, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, where ψ is the elasticity of intertemporal substitution (EIS). γ and ψ are required to satisfy $0 < \gamma, \psi$, and $\psi \neq 1$. For $\theta = 1$ the agent has standard CRRA preferences, while $\theta < 1$ indicates a preference for the early resolution of risk and $\theta > 1$ indicates a preference for late resolution. The general asset pricing equation to price any asset i with ex-dividend price $P_{i,t}$ and dividend $D_{i,t}$ is given by

$$E_t \left[M_{t+1} R_{i,t+1} \right] = 1 \tag{1.2}$$

where $R_{i,t+1} = \frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}$. For recursive preferences, the stochastic discount factor M_{t+1} is given by

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left(\frac{V_{t+1}}{\left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}}} \right)^{\frac{1}{\psi}-\gamma} . \tag{1.3}$$

Epstein and Zin (1989) show that the (unobserved) value of the aggregate wealth, W_t, can be expressed in terms of the value function,

$$W_t = \frac{V_t^{1-1/\psi}}{(1-\delta) C_t^{-1/\psi}} . \tag{1.4}$$

This expression in turn permits expressing M_{t+1} in terms of the gross return to the claim on aggregate consumption $R_{w,t+1}$,

$$M_{t+1} = \delta^{\theta} \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta-1} \tag{1.5}$$

where $R_{w,t+1} = \frac{W_{t+1}}{W_t - C_t}$. As equation (1.2) has to hold for all assets, it must also hold for the return of the aggregate consumption claim. Thus, $R_{w,t+1}$ is determined by the wealth-Euler equation

$$E_t \left[\delta^\theta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R^\theta_{w,t+1} \right] = 1. \tag{1.6}$$

Throughout the paper we consider a general setup for the specification of log consumption growth, Δc_{t+1}, that allows for long-run risk, x_t, and separate stochastic volatility processes, $\sigma_{c,t}$ and $\sigma_{x,t}$,

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_{c,t} \eta_{c,t+1} \\ x_{t+1} &= \rho x_t + \phi_x \sigma_{x,t} \eta_{x,t+1} \end{aligned} \tag{1.7}$$

where $\eta_{c,t+1}$ and $\eta_{x,t+1}$ are random shocks. In the remainder of the paper we consider variations of this setup that include different specifications for the stochastic volatility processes as well as additional state processes such as volatility of volatility or inflation.

Before we analyze the model we must answer two fundamental questions. First, does a solution for the model exist? Secondly, if a solution exists, how can we reliably compute it? To the best of our knowledge there are no closed-form solutions for the general model. So the common solution approach used in the finance literature is to log-linearize the model, see Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012a), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others. However, log-linearization misses by construction the influence of higher order dynamics; that is, the approach does not attempt to approximate nonlinear features of the exact solution. But what if these features matter qualitatively for the existence of solutions and quantitatively for equilibrium outcomes? Does log-linearization still deliver sufficiently accurate approximations of the exact solution?

We address these two critical issues in the next section of this paper. We first develop a formal existence criterion for the general model. Once we have established theoretical conditions, we can examine whether the log-linearized solution is in line with the formal existence results. For this task we need a different solution method that accurately accounts for higher-order dynamics and yields robust solutions. A convenient choice are projection methods that allow us to choose the approximation degree as well as the size of the approximation interval in order to be able to capture higher-order elements. While the projection methods require more computational effort, they are capable of correctly capturing higher-order features of the asset returns. For example Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012, p. 189) find that for a stochastic growth model with Epstein-Zin utility projection methods "provide a terrific level of accuracy with reasonable computational burden.” We compare the solutions obtained by log-linearization and projection and check whether the log-linearized solution provides reasonably accurate approximations.[2]

1.3 Existence and Computation of Solutions

Marinacci and Montrucchio (2010) and Hansen and Scheinkman (2012) consider the existence of solutions for the fixed-point equation for the value function for general process specifications. Applying these results to the Bansal-Yaron model has proven delicate. For example, de Groot (2015) considers the existence of solutions for growth economies with stochastic volatility under CRRA preferences, and finds that existence is a complex nonlinear function of the process.

To prove existence, we sidestep the challenge of proving a general result, but instead provide a simple relative result. We show that if the model has a solution for CRRA preferences ($\theta = 1$), then it has a solution when investors have a preference for early resolution of risk ($\theta < 1$), which includes most models in the literature. Interestingly, if investors have a preference for late resolution of uncertainty ($\theta > 1$), the implication is reversed.

This relative result allows us to leverage the extensive literature proving existence for CRRA utility for growth economies. The initial contribution of Burnside (1998) provides both a closed-form solution and characterization for existence when log-consumption growth follows a simple AR(1) process with Gaussian shocks. This result has been extended in various ways. Bidarkota and McCulloch (2003) and Tsionas (2003) generalize the result by relaxing the assumption of normal shocks to any stable shock distribution and to shocks with well-defined moment generating functions, respectively. Collard, Féve, and Ghattassi (2006) show how to generalize Burnside (1998) to the case of habit formation. Calin, Chen, Cosimano, and Himonas (2005) derive closed-form solutions for asset pricing models with one state variable as long as the utility function and the price-dividend function are analytic. Chen, Cosimano, and Himonas (2008) use this method to analyze existence of solution in the habit model of Campbell and Cochrane (1999) and show how to generalize the approach to multi-dimensional state spaces. Most directly relevant to our application, de Groot (2015) shows how to generalize the result to processes that feature stochastic volatility. In an online appendix, deGroot also provides a closed-form solution for both long-run risk and stochastic volatility, as in the specification of Bansal and Yaron (2004). The results can be generalized further to specifications featuring, for example, volatility of volatility or inflation.

[2] In Chapter 4 of this thesis we provide a detailed description of how projection methods can be used to solve asset pricing models with Epstein-Zin preferences. In Appendix 1.B we show an example that demonstrates the high accuracy of projection methods.

1.3.1 Existence for Epstein-Zin Utility

For the proof of a formal existence theorem for the general model, we first state a special fixed-point result. Subsequently, we present an existence theorem. Appendix 1.A contains the proofs of all formal statements.

1.3.1.1 A Fixed-Point Result

Marinacci and Montrucchio (2010) apply Tarski's Fixed-Point Theorem, Tarski (1955) to establish the existence of solutions to general nonlinear stochastic equations which encompass, as special cases, many of those arising in stochastic dynamic programming. Here we use a similar fixed-point argument, such as in the proof of Proposition 1 in Marinacci and Montrucchio (2010, Section 4), in a key step towards proving the existence of solutions to the asset pricing equation for Epstein-Zin utility.

Let s_t be a real vector-valued Markov process with elements in $\mathcal{S} \subset \mathbb{R}^l$, $l \geq 1$, with conditional probability density $p(s'|s)$. Let $\mathcal{V}$ be the set of all Lebesgue-measurable functions $f : \mathcal{S} \to \mathbb{R}_+$, such that

$$\int f(s')^{\max(1,\theta)} p(s'|s) < \infty.$$

This set is the space of all candidate solutions to the fixed-point problems addressed in the lemmata below. We write $f \geq g$ if $f(s) \geq g(s)$ for almost all s. This introduces a partial order on $\mathcal{V}$. With this partial order, for any given $g^* \in \mathcal{V}$, the interval $[0, g^*] \equiv \{f \in \mathcal{V} : 0 \leq f \leq g^*\}$ is a complete lattice.

Now consider a functional $T : \mathcal{V} \to \mathcal{V}$. The functional T is monotone, or order-preserving, iff $f \leq g$ implies $Tf \leq Tg$ for any pair $f, g \in \mathcal{V}$. Further suppose that T maps $[0, g^*]$ to itself for some $g^* \in \mathcal{V}$, so $T([0, g^*]) \subseteq [0, g^*]$. Then the Tarski Fixed-Point Theorem, Tarski (1955), implies that T has a fixed point in $[0, g^*]$; in fact, the set of fixed points is also a complete lattice. This theorem implies the following lemma.

Lemma 1. *Let $T, U : \mathcal{V} \to \mathcal{V}$, such that for any pair $f, g \in \mathcal{V}$ with $f \leq g$ it holds that*

$$Tf \leq Tg \leq Ug.$$

Further suppose g^* is a fixed point of U. Then T has a fixed point in $[0, g^*]$.

We also need the following lemma.

Lemma 2. *Let $\beta \in [0, 1)$, $\lambda \in \mathbb{R}$, and $0 \neq \theta \leq 1$. Furthermore, let $C, f \in \mathcal{V}$. Let T be the operator*

$$Tf = (1 - \beta)C^\lambda + \beta \left[E(f^\theta|s)\right]^{1/\theta}$$

and U be the operator

$$Uf = (1 - \beta)C^\lambda + \beta E(f|s).$$

Then T and U preserve $\mathcal{V}$. In addition,

(A) Let $0 \neq \theta \leq 1$. Then for any $f, g \in \mathcal{V}$ with $f \leq g$ implies $Tf \leq Tg \leq Ug$.

(B) Let $\theta > 1$. Then for any $f, g \in \mathcal{V}$ with $f \leq g$ implies $Uf \leq Ug \leq Tg$.

Applying Lemma 1 to the operators in Lemma 2 leads to our final conclusion in this section.

Lemma 3. *For the two operators in Lemma 2(A), if U has a fixed point, so does T. For the two operators in Lemma 2(B), if T has a fixed point, so does U.*

This lemma enables us to obtain an existence result for the model with Epstein-Zin (EZ) preferences.

1.3.1.2 CRRA vs. EZ

Recall the value function recursion (1.1) for Epstein-Zin utility,

$$V_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

with $0 < \gamma, \psi \neq 1$, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. The value $\theta = 1$ yields CRRA utility as a special case. Define $\lambda = 1 - \frac{1}{\psi}$ and $\hat{V} = V^{\lambda}$ to obtain

$$\begin{aligned} V_t^{\lambda} &= (1-\delta) C_t^{\lambda} + \delta \left[E_t \left((V_{t+1}^{\lambda})^{\frac{1-\gamma}{\lambda}} \right) \right]^{1/\theta} \\ \iff \quad \hat{V}_t &= (1-\delta) C_t^{\lambda} + \delta \left[E_t \left(\hat{V}_{t+1}^{\theta} \right) \right]^{1/\theta}. \end{aligned} \tag{1.8}$$

The following theorem relates solutions for the model with CRRA utility ($\theta = 1$) to solutions for the general model with $\theta \neq 1$.

Theorem 1. *Let $0 < \psi \neq 1$ be given. Suppose consumption C is a positive function of a real vector-valued Markov process.*

(A) If the asset pricing model characterized by equations (1.1)–(1.6) *has a solution for CRRA utility, $\gamma = \frac{1}{\psi}$, then it also has a solution for Epstein-Zin utility with $0 \neq \theta < 1$; that is, for $1 \neq \gamma > \frac{1}{\psi}$ if $\psi > 1$ and $1 \neq \gamma < \frac{1}{\psi}$ if $\psi < 1$.*

(B) If the asset pricing model characterized by equations (1.1)–(1.6) *has a solution for Epstein-Zin utility with $\theta > 1$, that is, for $\gamma < \frac{1}{\psi}$ if $\psi > 1$ and $\gamma > \frac{1}{\psi}$ if $\psi < 1$, then it also has a solution for CRRA utility with $\gamma = \frac{1}{\psi}$.*

Theorem 1(A) enables us to use existence results for the CRRA case that can be derived for various state-process specifications (see, among others, Burnside (1998) or de Groot (2015)) to determine regions for the parameters ψ and γ for which a solution also exists for EZ utility. The contrapositive of Theorem 1(B) enables us to use the CRRA non-existence results of this literature to determine regions for the parameters ψ and γ for which no solution exists for EZ utility. Since under Epstein-Zin utility, by equation (1.4),

$$W_t = \frac{\hat{V}}{(1-\delta)C_t^{-1/\psi}},$$

the bound on $\hat{V}$ translate immediately to a bound on wealth, and the wealth-consumption ratio. We will see in our numerical results that this bound is satisfied by the numerical approximations.

Theorem 1 allows for a general consumption process. Next we consider the special state-process specification of Bansal and Yaron (2004). We consider this special specification since it has received much attention in the finance literature. We first analyze the model with CRRA utility and subsequently analyze the implications for general Epstein-Zin preferences using the statements of Theorem 1.

1.3.2 Existence in the Long-Run Risk Model: CRRA Utility

Bansal and Yaron (2004) use the state processes as in equation (1.7) with Gaussian shocks and assume that the stochastic volatility in the economy is captured by a single volatility process ($\sigma_{c,t} = \sigma_{x,t} = \sigma_t$):

$$\sigma_{t+1}^2 = \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \phi_\sigma \omega_{t+1}, \tag{1.9}$$

with $\eta_{c,t+1}, \eta_{x,t+1}, \omega_{t+1} \sim$ *i.i.d.* $N(0,1)$.[3] The following theorem states a formal condition that ensures a finite wealth-consumption ratio and hence the existence of a solution for the model with CRRA utility ($\theta = 1$). Appendix 1.A outlines a proof which closely follows the arguments in de Groot (2015).

Theorem 2. *There exists a solution to model* (1.1)–(1.7) *with* $\theta = 1$ *and a single volatility process as specified in equation* (1.9) *if and only if*

$$\delta \exp\left(\underbrace{B}_{Constant} + \underbrace{B_c\bar{\sigma}^2}_{Consumption\ Shock} + \underbrace{B_x\phi_x^2}_{LRR\ Shock} + \underbrace{B_\sigma\phi_\sigma^2}_{SV\ Shock} \right) < 1, \tag{1.10}$$

[3] The assumption of normal shocks is not necessary in general for the derivation of closed-form solutions; it suffices that the moment generating function of the shocks exists. For example, de Groot (2015) also provides solutions for a truncated normal and a gamma distribution. These distributions offer the great advantage that the variance process remains positive. However, most of the research following the seminal work by Bansal and Yaron (2004) adopts the normal assumption, which motivates our focus on the existence of solutions for this model class.

with the following coefficients $\mathcal{B} = (B, B_c, B_x, B_\sigma)$,

$$\begin{aligned}
B &= \mu_c \left(1 - \frac{1}{\psi}\right), \\
B_c &= 0.5 \left(1 - \frac{1}{\psi}\right)^2, \\
B_x &= 0.5 \left(\frac{1 - \frac{1}{\psi}}{1 - \rho}\right)^2 \bar{\sigma}^2, \\
B_\sigma &= \frac{1}{8} \left(\left(\frac{\left(1 - \frac{1}{\psi}\right)^2 \phi_x^2}{(1-\rho)(1-\nu)} \right)^2 + 2 \left(\frac{\left(1 - \frac{1}{\psi}\right)^2 \phi_x}{(1-\rho)(1-\nu)} \right)^2 + \left(\frac{\left(1 - \frac{1}{\psi}\right)^2}{(1-\nu)} \right)^2 \right),
\end{aligned} \tag{1.11}$$

which only depend on the parameters of the state processes and the intertemporal elasticity of substitution, ψ.

Expression (1.10) shows that the existence of solutions depends on the size of the subjective discount factor δ and a constant part B. In addition, each shock in the model, $\eta_{c,t+1}, \eta_{x,t+1}$ and $\eta_{\sigma,t+1}$, adds a new term to the existence requirement.[4] The presence of each type of shock makes the existence requirement more demanding since the three coefficients B_c, B_x, and B_σ are all positive. In the following, we decompose expression (1.10) to analyze the influence of the three different factors on the existence of solutions. Rewriting the inequality yields

$$B + B_c \bar{\sigma}^2 + B_x \phi_x^2 + B_\sigma \phi_\sigma^2 < -\ln \delta. \tag{1.12}$$

For the baseline calibration of Bansal and Yaron (2004) with a value of $\delta = 0.998$, the sum over the four components on the left-hand side must be smaller than 0.002. For a larger discount factor δ, as, for example, in the study of Schorfheide, Song, and Yaron (2014) with a value of 0.9996, the condition becomes more stringent with a right-hand side of only 0.0004.

Observe that $\frac{\partial B_x \phi_x^2}{\partial \rho} > 0$, $\frac{\partial B_x \phi_x^2}{\partial \phi_x} > 0$, $\frac{\partial B_\sigma \phi_\sigma^2}{\partial \nu} > 0$ and $\frac{\partial B_\sigma \phi_\sigma^2}{\partial \phi_\sigma} > 0$.[5] Thus the higher the volatility and persistence of the state processes, the more stringent becomes the condition for existence. Table 1.1 reports the magnitudes of the four terms on the left-hand side of (1.12) for two different parameterizations. In particular, we provide values for a conservative calibration for the long-run risk process and the stochastic volatility channel and a calibration that takes the more extreme values found in the literature. (Compare Table 1.2 in Section 1.4 for the parameter values from six recent studies in the finance literature.)

[4] Note that in this model specification stochastic volatility influences not only shocks to consumption but also shocks to long-run risk. In a more parsimonious setup, with stochastic volatility only entering the shocks to consumption, where the long-run risk factor is a standard AR(1) process ($\sigma_{x,t} = 1$, $\forall t$) the coefficients simplify to $B_x = 0.5 \left(\frac{1-\frac{1}{\psi}}{1-\rho}\right)^2$ and $B_\sigma = \frac{1}{8} \frac{\left(1-\frac{1}{\psi}\right)^4}{(1-\nu)^2}$ and so there is no interaction between the separate terms.

[5] Note that, since stochastic volatility also affects the long-run risk factor, it holds that $\frac{\partial B_\sigma \phi_\sigma^2}{\partial \rho} > 0$ and $\frac{\partial B_\sigma \phi_\sigma^2}{\partial \phi_x} > 0$, making the conditions for existence more stringent as ρ and ϕ_x increase.

Table 1.1: Existence in the Long-Run Risk Model of Bansal and Yaron (2004)

Conservative Estimates

	B	$B_c\bar{\sigma}^2$	$B_x\phi_x^2$	$B_\sigma\phi_\sigma^2$	Sum
$\psi = 2$	0.00075	6.5e-6	1.5e-5	1.2e-10	0.00077
$\psi = 1.5$	0.00050	2.9e-6	6.7e-6	2.4e-11	0.00051
$\psi = 0.5$	-0.00150	2.6e-5	6.0e-5	1.9e-9	-0.00141
$\psi = 0.2$	-0.00600	4.1e-4	9.6e-4	4.9e-7	-0.00462

High Estimates

	B	$B_c\bar{\sigma}^2$	$B_x\phi_x^2$	$B_\sigma\phi_\sigma^2$	Sum
$\psi = 2$	0.00075	7.6e-6	0.00030	4.9e-6	0.00106
$\psi = 1.5$	0.00050	3.4e-6	0.00013	9.7e-7	0.00064
$\psi = 0.5$	-0.00150	3.0e-5	0.00120	7.8e-5	-0.00019
$\psi = 0.2$	-0.00600	4.9e-4	0.01923	0.0201	0.03381

The table displays values for the four terms in condition (1.12) *which determine the existence of solutions in the model of Bansal and Yaron (2004) for two sets of parameter calibrations. The conservative parameters are given by* $\bar{\sigma} = 0.0072, \rho = 0.975, \phi_x = 0.038, \nu = 0.956, \phi_\sigma = 2.3e\text{-}6$. *The high estimates are* $\bar{\sigma} = 0.0078, \rho = 0.993, \phi_x = 0.044, \nu = 0.999, \phi_\sigma = 2.8e\text{-}6$. *For both cases we use* $\mu_c = 0.0015$.

For a monthly time interval $\mu_c \approx 0.0015$ and the long-run risk literature argues in favor of $\psi \approx 1.5$. These estimates yield a constant term of $B = 0.0005$. For $\psi > 1$, the constant B increases in ψ making the existence condition more stringent as ψ increases. For example, for a value of $\psi = 2$, B becomes 0.0075. Among others, Campbell (1996), Attanasio and Weber (1995) and Yogo (2004) argue for an elasticity of substitution below one. In that case, the constant B becomes negative and hence relaxes the existence condition.

For the conservative parameter range, Table 1.1 shows that the sum of the four terms on the left-hand side of condition (1.12) is always (clearly) below 1e-3. Therefore, as long as $\delta < 0.999$, condition (1.12) easily holds and the model has a solution. For the high parameters estimates the influence of the consumption shock $B_c\bar{\sigma}^2$ is still very small. The influence of the long-run risk process, $B_x\phi_x^2$, strongly increases and assumes values between 0.00013 and 0.02 depending on the EIS ψ. In light of the condition (1.12), we observe that adding long-run risk to the model can have strong effects on the existence of solutions. The influence of the stochastic volatility shock $B_\sigma\phi_\sigma^2$ remains rather insignificant for an EIS larger than one, but increases strongly as an EIS less than one decreases further. In particular, we observe that for $\psi = 0.2$ the model has no solution for $\delta > e^{-0.03381} \approx 0.9668$.

This completes our discussion of existence in the long-run risk model with CRRA preferences. We now combine the insights from Theorems 1 and 2 to analyze the existence of solutions for the model with Epstein-Zin utility.

1.3.3 Existence in the Long-Run Risk Model: EZ Utility

While there is much debate[6] in the economics and finance literature whether the elasticity of substitution is larger or smaller than one, there appears to be widespread agreement on parameters that satisfy $\gamma > 1/\psi$. Thus, we now restrict attention to models with such preferences. Recall from Theorem 1(A) that, if the model has a solution for CRRA preferences with $\psi > 1$, it also has a solution for recursive preferences with $\gamma > 1/\psi$. And so, for cases such as $\psi = 1.5$ and $\psi = 2$, which we consider in the following, the model with recursive preference with $\gamma > 1/\psi$ has a solution for any exogenous consumption specification satisfying condition (1.12) in Theorem 2. On the contrary, the contrapositive of Theorem 1(B) shows that, if there is no solution for CRRA preferences with $\psi < 1$, then the model with $\gamma > 1/\psi$ also cannot have a solution. And so, for cases such as $\psi = 0.2$ and $\psi = 0.5$, which we consider in the following as well, the model with recursive preference with $\gamma > 1/\psi$ does not have a solution for any exogenous consumption specification violating condition (1.12) in Theorem 2.

In the following we illustrate these implications of the two parts of Theorem 1. We first analyze the effects of long-run risk and stochastic volatility on the existence of solutions separately. Subsequently, we examine the existence properties in the full calibrated long-run risk model of Bansal and Yaron (2004).

1.3.3.1 Long-Run Risk without Stochastic Volatility

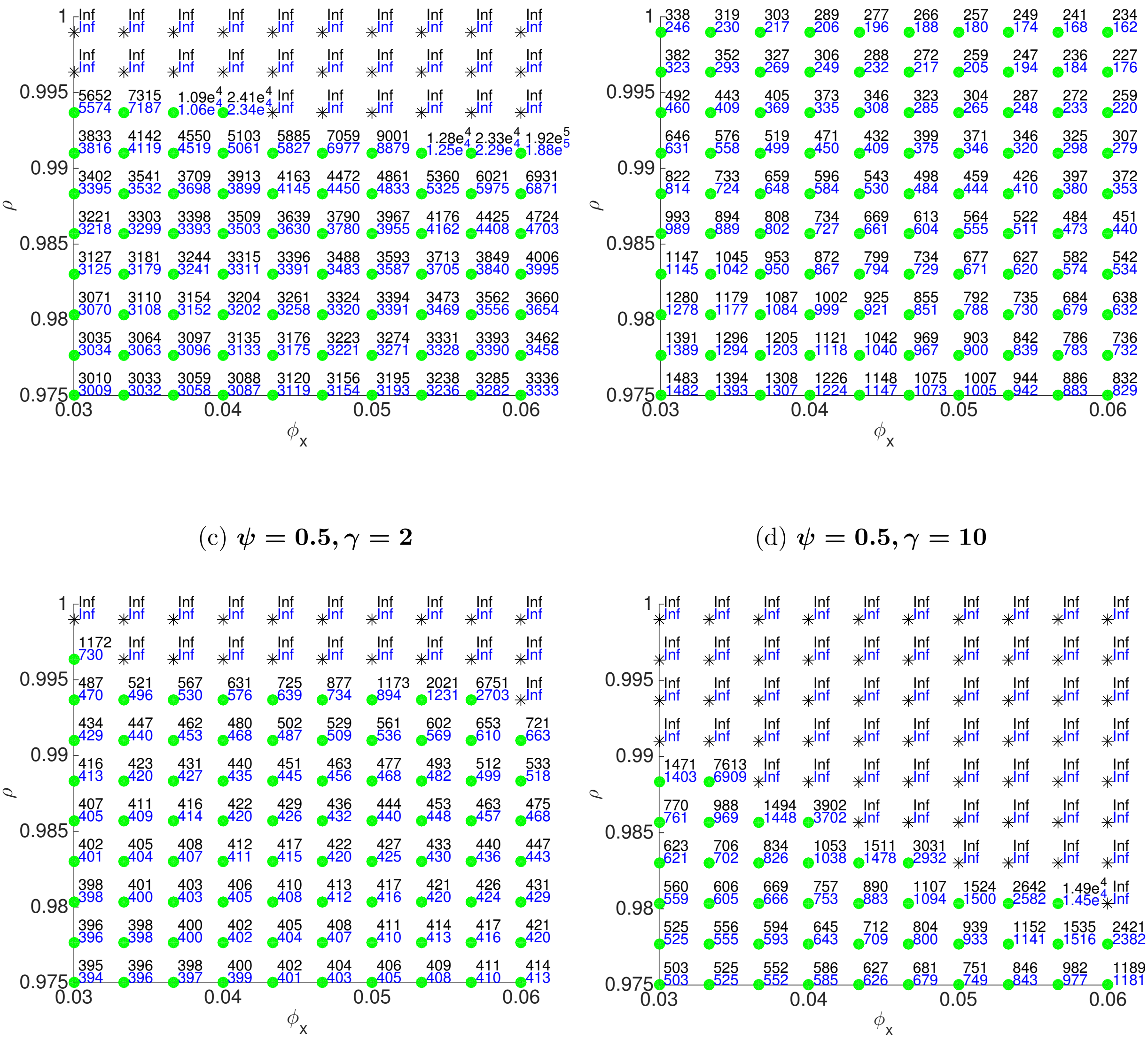
To obtain an impression of the isolated effect of long-run risk on the existence of solution with recursive utility, we "shut off" stochastic volatility by setting $\sigma_{c,t} = \sigma_{x,t} = \bar{\sigma}$. Since there is a large debate in the asset pricing literature about the right calibration of the persistence ρ and volatility ϕ_x of the long-run risk process (see for example Bansal, Kiku, and Yaron (2012a), Beeler and Campbell (2012), Schorfheide, Song, and Yaron (2014) and Bollerslev, Xu, and Zhou (2015)), we provide solutions for a range of parameters. Figure 1.1 shows convergence properties as well as the mean wealth-consumption ratio obtained by the log-linearization and the projection method for a 10×10 grid of values for ϕ_x and ρ. For the case of CRRA utility there are closed-form solutions for the model (see de Groot (2015)) and Theorem 2 shows the formal existence condition. For the case of recursive utility we compute highly accurate

[6] Table 1.2 in Section 1.4 displays EIS values from six recent studies in the asset pricing literature, namely those of Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012a), Bollerslev, Xu, and Zhou (2015), Schorfheide, Song, and Yaron (2014) and Bansal and Shaliastovich (2013). These studies estimate the long-run risk model by trying to match (asset pricing) moments and obtain values between 1.5 and 2. On the contrary, Yogo (2004) provides estimates below 0.2 using a linearized Euler equation and matching the interest rate. Attanasio and Weber (1995) reports estimates of 0.67 and smaller depending on the data set.

Figure 1.1: Influence of Long-Run Risk on Existence and Higher-Order Dynamics

(a) $\boldsymbol{\psi = 2, \gamma = 0.5}$ (b) $\boldsymbol{\psi = 2, \gamma = 10}$

(c) $\boldsymbol{\psi = 0.5, \gamma = 2}$ (d) $\boldsymbol{\psi = 0.5, \gamma = 10}$



The graph shows the convergence properties as well as the mean wealth-consumption ratio for model (1.7) with constant volatility $\sigma_{c,t} = \sigma_{x,t} = \bar{\sigma}$ and i.i.d. normal shocks $\eta_{c,t+1}$ and $\eta_{x,t+1}$. The results are reported for a range of persistence parameters ρ and volatility parameters ϕ_x. Panels (a) and (c) depict the cases of CRRA utility with $\psi = 2$ and $\psi = 0.5$ respectively, while panels (b) and (d) depict the corresponding cases with EZ utility and $\gamma = 10$. Green circles denote convergence of both, the projection and the log-linearization approach. In the case of CRRA preferences the formal existence condition (1.10) is also satisfied. Black stars denote cases in which both methods don't converge and the model also doesn't have a solution in the case of CRRA preferences. Black numbers show the mean wealth-consumption ratio obtained by the projection approach and blue numbers show the values obtained by the log-linearization. The remaining model parameters are given by $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0078$.

solutions using the projection approach.[7] A (green) circle indicates for CRRA utility that the convergence condition of Theorem 2 is satisfied; for EZ utility the circle indicates that both methods produce a solution. A (black) star indicates for CRRA utility that no solution exists and for EZ utility that the projection method does not converge. The lower (blue) values in the figure show the mean wealth-consumption ratio for the log-linearization and the upper (black) values for the projection approach. The entry "Inf" indicates that a method did not find a solution.[8]

Panel (a) of Figure 1.1 shows that the model has a solution for CRRA preferences with $\psi = 2$ for sufficiently low values of the volatility parameter and the persistence. In line with Theorem 2, the convergence condition becomes more stringent the larger the persistence, ρ, or the larger the volatility, ϕ_x, and there is no convergence for high-volatility high-persistence combinations. Panel (b) displays the corresponding results for recursive utility with $\psi = 2$, $\gamma = 10$. We find that there is convergence for all parameter combinations on the selected grid. This finding is in line with Theorem 1(A) that, if there exists a solution for the CRRA utility case with $\psi > 1$, there also exists a solution for the model with recursive preferences and $\gamma > 1/\psi$. Put differently, for $\psi > 1$ increasing the risk aversion parameter γ leads to a less stringent existence condition.

Panel (c) of Figure 1.1 depicts the case of CRRA preferences but with $\psi = 0.5$. Again the model is well behaved in the region of low volatility and low persistence region and the convergence condition becomes more stringent the higher the persistence ρ or the higher the volatility ϕ_x. However, for the corresponding case with recursive preferences and $\gamma = 10$ (Panel (d)), there is no convergence for a much larger set of parameters. Hence, in the case of $\psi < 1$ increasing the coefficient γ makes the existence condition more stringent. This finding is consistent with Theorem 1(B) which shows that the existence condition is more demanding for $\psi < 1$ and $\gamma > 1/\psi$ compared to the respective CRRA case.

---

[7] A formal analysis of the accuracy of the projection approach is conducted in Appendix 1.B. To compute accurate solutions with the projection method we increase the approximation interval and the polynomial approximation degree until the solutions no longer change and the polynomial coefficients for the highest degree polynomial are close to zero. By this approach we make sure that we capture the higher-order dynamics introduced by the tails of the state processes. For the case with CRRA utility we obtain the same solution as the closed-form expressions derived by de Groot (2015) (up to some tiny error). For the cases where there don't exist closed-form solutions we double-checked the accuracy of our computations by using the discretization technique of Tauchen and Hussey (1991) with a very large number of discretization nodes.

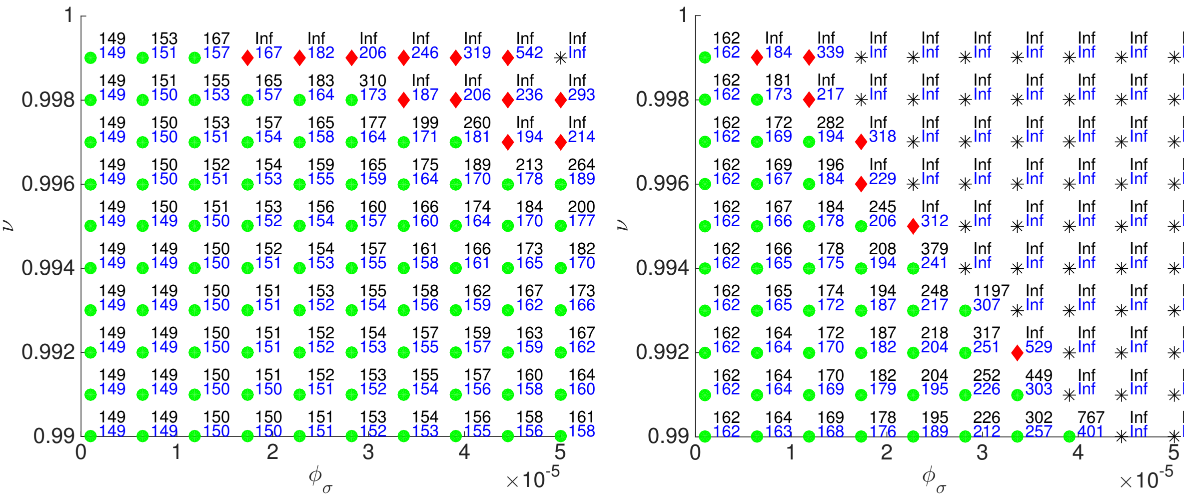
[8] Both solution methods ultimately require us to solve a nonlinear system of equations. If the solver cannot solve the system for the log-linearization approach, then, as a robustness check, we attempt to find a solution by setting up a grid of 1000 starting points for the linearization constant. Only if the solver still cannot find a solution, do we report "Inf" for the log-linearization method. If the solver cannot solve the system for the projection method, then we first attempt to compute a solution for a very small state space and a small degree of the approximating polynomial. Subsequently we increase the state space and the polynomial degree. As initial guesses we use solutions from model specifications where we found a solution. While a complete failure of many repeated attempts with the projection method to find a solution are not a proof of non-existence, they give us a high degree of confidence that indeed no solution exists. Also, in the case of CRRA preferences, this approach yields exactly the same convergence results as obtained by the formal Theorem 2.

1.3.3.2 Stochastic Volatility without Long-Run Risk

Figure 1.2 shows the results for the model with stochastic volatility but without long-run risk ($x_t = 0\ \forall t$). Recall from Table 1.1 that the influence of the stochastic volatility channel is especially strong for low values of the EIS. Therefore, Panel (a) displays solutions for the CRRA case with an EIS of $\psi = 0.2$ and Panel (b) for the corresponding EZ case with $\gamma = 10$ for a 10×10 grid of values for the parameters ϕ_σ and ν.

Figure 1.2: Influence of Stochastic Volatility on Existence and Higher-Order Dynamics

(a) $\boldsymbol{\psi = 0.2, \gamma = 5}$ (b) $\boldsymbol{\psi = 0.2, \gamma = 10}$



The graph shows the convergence properties as well as the mean wealth-consumption ratio for model (1.7) with no long-run risk ($x_t = 0\ \forall t$) and a single stochastic volatility process given by equation (1.9). The results are reported for a range of persistence parameters ν and volatility parameters ϕ_σ. Panel (a) depicts the cases of CRRA utility with $\psi = 0.2$ while panels (b) depicts the corresponding cases with EZ utility and $\gamma = 10$. Green circles denote convergence of both, the projection and the log-linearization approach. In the case of CRRA preferences the formal existence condition (1.10) is also satisfied. Black stars denote cases in which both methods don't converge and the model also doesn't have a solution in the case of CRRA preferences. Red diamonds denote the cases in which the model doesn't have a solution, but the log-linearization gives a finite wealth-consumption ratio. Black numbers show the mean wealth-consumption ratio obtained by the projection approach and blue numbers show the values obtained by the log-linearization. The remaining model parameters are given by $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0072$.

We now observe a new phenomenon which we represent by (red) diamonds in the figure. For the CRRA case, the diamonds depict parameter combinations for which the model does not have a solution, the condition in Theorem 2 is violated, but the log-linearization approach yields a finite wealth-consumption ratio and incorrectly indicates existence. Simply put, the log-linearization approach delivers a model solution even though the model does not have a(n exact) solution. On the contrary, the projection method correctly indicates nonexistence for all these cases. For the specification with Epstein-Zin utility in panel (b), the diamonds indicate parameter combinations for which the projection method indicates nonexistence while the log-linearization approach indicates existence.

What is the reason for the failure of the log-linearization approach? Whenever the model does not have a solution, the wealth-consumption ratio is in fact infinite. As we see from the reported values for the mean wealth-consumption ratio in Figure 1.2, the log-linearization systematically underestimates the wealth-consumption ratio. This underestimation becomes especially strong in the regions close to non-existence leading to fundamentally wrong model outcomes in this parameter region. In addition to this qualitative effect, we also observe a strongly related quantitative effect. For fixed persistence ν of the volatility process, the degree of underestimation increases in the volatility parameter ϕ_σ. That is, the numerical error of the log-linearization result increases in ϕ_σ (until it eventually becomes infinitely large).

1.3.3.3 The Long-run Risk Model Calibration of Bansal and Yaron (2004)

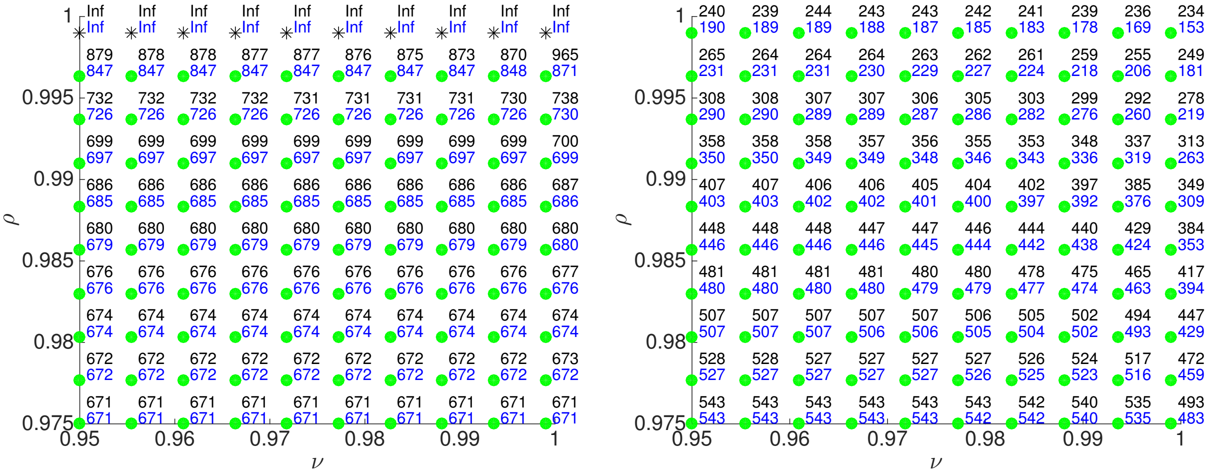
In the third and final step of our numerical existence analysis, we show the simultaneous effects of long-run risk and stochastic volatility on the existence of solutions for the long-run risk model of Bansal and Yaron (2004). Figure 1.3 depicts the convergence properties and the mean wealth-consumption ratios for a grid of values for the persistence parameters of the long run risk process, ρ, and the stochastic volatility, ν.

Panel (a) shows results for CRRA utility, while Panel (b) shows results for the the utility parameters of Bansal and Yaron (2004). With the exception of models with very high values of the persistence ρ of the long-run risk factor and CRRA utility, the models have a solution. In accordance with Theorem 1(A), increasing the risk aversion to $\gamma = 10$ increases the region of convergence because $\psi > 1$. In line with the results reported in Table 1.1 in Section 1.3.2, the stochastic volatility channel does not significantly affect the existence region (the non-existence region does not grow (significantly) with ν) due to its relatively low volatility in the calibration of Bansal and Yaron (2004). Put differently, the additional feature of stochastic volatility in long run risk models has a negligible influence on the qualitative existence issue of solutions. However, the stochastic volatility does have a strong quantitative effect on the approximation errors of the log-linear solution of the model, particularly in Panel (b) which shows the results for the utility parameters of Bansal and Yaron (2004), $\psi = 1.5$, $\gamma = 10$. We observe that both the absolute and the relative difference between the log-linearized and the true model solution increases substantially with both persistence parameters ρ and ν of the long run risk and the stochastic volatility, respectively. Apparently, adding another state process to the model introduces new non-linearities which depend strongly on the persistence of the process.

We emphasize that our analysis of models with very high values for the persistence parameters is not an artificial exercise. In fact, as we report in Table 1.2 in the next section, recent work on asset pricing models regularly uses highly persistent processes for the exogenous model inputs. The stochastic volatility and the long-run risk process in Bansal and Yaron (2004), the inflation processes in Bansal and Shaliastovich (2013) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) or the different volatility processes in Schorfheide, Song, and Yaron

Figure 1.3: Existence and Higher-Order Dynamics in the Long-Run Risk Model

(a) $\boldsymbol{\psi = 1.5, \gamma = 2/3}$ (b) $\boldsymbol{\psi = 1.5, \gamma = 10}$



The graph shows the convergence properties as well as the mean wealth-consumption ratio for the long-run risk model of Bansal and Yaron (2004). The results are reported for a range of persistence parameters of the long run risk process ρ and the stochastic volatility process ν. Panels (a) depicts the cases of CRRA utility with $\psi = 1.5$, while panel (b) depicts the corresponding cases with EZ utility and $\gamma = 10$. Green circles denote convergence of both, the projection and the log-linearization approach. In the case of CRRA preferences the formal existence condition (1.10) is also satisfied. Black stars denote cases in which both methods don't converge and the model also doesn't have a solution in the case of CRRA preferences. Black numbers show the mean wealth-consumption ratio obtained by the projection approach and blue numbers show the values obtained by the log-linearization. The remaining model parameters are given by $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0078, \phi_x = 0.044, \phi_\sigma = 2.3\text{e-}6$.

(2014) are a few examples of such processes. In all those papers, log-linearization techniques have been used to analyze equilibrium quantities. But as we have demonstrated above, solving highly persistent models using log-linearization can introduce large approximation errors in the mean wealth-consumption ratio. Naturally now the question arises whether these errors also matter for the model predictions of economically relevant quantities such as, for example, the equity premium, the risk free rate, or return volatilities; or whether perhaps these errors have only small effects on these quantities and so log-linearization remains a reliable solution approach for such model predictions. We answer this question for a number of prominent asset pricing models in the next section.

1.4 Higher-Order Dynamics

In this section we compare the implications of the solutions of the log-linearization approach and the projection method for a number of economically relevant quantities. Specifically, we perform this comparison for six different studies from the asset pricing literature on long run risk. The six models are the seminal long-run risk model of Bansal and Yaron (2004), the re-calibrated version of the model by Bansal, Kiku, and Yaron (2012a), the extensive estimation study of Schorfheide, Song, and Yaron (2014), the volatility-of-volatility models of Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Xu, and Zhou (2015), and the two studies study of real and nominal bonds of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013). Common to all these studies is the methodological attempt to match several key statistics on financial markets such as the high equity premium, a low risk-free rate, volatile stock prices, real and nominal bond prices, the volatility premium or patterns in return predictability. Obviously, in order to determine a reasonable calibration of the model it is essential to solve the model without significant errors in the approximation of those key statistics since such errors could potentially bias the calibration.

In the previous section, we have seen that the log-linearization approach produces sizable approximation errors for the mean wealth-consumption ratio in the long-run risk model of Bansal and Yaron (2004). Now we show that these errors carry forward to substantial errors in the first and second moments of asset returns. In fact, we demonstrate that making use of the log-linearization approach has a strong impact on the financial market statistics implied by the models.

1.4.1 Six Model Specifications

The models share the same basic model setup (1.7) augmented with a process for log dividend growth Δd_{t+1} that is potentially correlated with consumption,

$$\begin{aligned}\Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_{c,t} \eta_{c,t+1} \\ x_{t+1} &= \rho x_t + \phi_x \sigma_{x,t} \eta_{x,t+1} \\ \Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma_{d,t} \eta_{d,t+1} + \phi_{d,c} \sigma_{c,t} \eta_{c,t+1} \\ &\eta_{c,t+1}, \eta_{x,t+1}, \eta_{d,t+1} \sim i.i.d.\ N(0,1).\end{aligned} \tag{1.13}$$

Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a) assume that there is a single volatility process that drives uncertainty in the economy $\sigma_{c,t} = \sigma_{x,t} = \sigma_{d,t} = \sigma_t$ with

$$\sigma_{t+1}^2 = \bar{\sigma}^2(1-\nu) + \nu \sigma_t^2 + \phi_\sigma \omega_{t+1} \qquad \omega_{t+1} \sim i.i.d.\ N(0,1). \tag{1.14}$$

Schorfheide, Song, and Yaron (2014) relax this assumption by allowing for three separate volatility processes. The two volatility processes for consumption growth and the long-run risk factor are required to account for the weak correlation between the risk-free rate and consumption growth. As shown in their estimation study, the volatility dynamics of dividends differs significantly from the other two processes. Therefore, a third process is required to model the stochastic volatility of dividends. Schorfheide, Song, and Yaron (2014) assume that the logarithm of the volatility process is normal to ensure that the standard deviation of the shocks remains positive,

$$\begin{aligned} \sigma_{i,t} &= \varphi_i \bar{\sigma} \exp(h_{i,t}) \\ h_{i,t+1} &= \nu_i h_{i,t} + \sigma_{h_i} \sqrt{1-\nu_i^2}\omega_{i,t+1}, \qquad i \in \{c, x, d\} \\ \omega_{i,t+1} &\sim i.i.d.\ N(0,1). \end{aligned} \tag{1.15}$$

In order to derive analytical solutions for the log-linearization coefficients that are needed for their estimation study, Schorfheide, Song, and Yaron (2014) use a linear approximation of the volatility dynamics that follows Gaussian dynamics,

$$\sigma_{i,t}^2 \approx 2(\varphi_i \bar{\sigma})^2 h_{i,t} + (\varphi_i \bar{\sigma})^2 \tag{1.16}$$

which in turn yields

$$\sigma_{i,t+1}^2 = \bar{\sigma}_i^2 (1-\nu_i) + \nu_i \sigma_{i,t}^2 + \phi_{\sigma_i} \omega_{i,t+1}$$

with $\phi_{\sigma_i} = 2\bar{\sigma}_i^2 \sigma_{h_i} \sqrt{1-\nu_i^2}$ and $\bar{\sigma}_i = \varphi_i \bar{\sigma}$.[9]

The fourth model stems from the estimation study of Bollerslev, Xu, and Zhou (2015). In a standard long-run risk model with stochastic volatility many long-standing puzzling behaviors on financial markets such as a high equity risk premium together with a low risk-free rate, volatile price dynamics or predictability of stock returns can be explained. However, the most recent research has gone one step further by showing that the standard model is not able to generate a time-varying variance risk premium that has predictive power for stock returns. Fortunately, the literature has also suggested a possible solution for this puzzle by including time-varying volatility of volatility (vol-of-vol) to the model, see, for example, Bollerslev, Tauchen, and Zhou (2009), Tauchen (2011), Drechsler and Yaron (2011), Bollerslev, Xu, and Zhou (2015) or Dew-Becker, Giglio, Le, and Rodriguez (2015). Bollerslev, Xu, and Zhou (2015) consider a slight variation of the long-run risk factor compared to the baseline model (1.13)

[9] We proceed in the same way as Schorfheide, Song, and Yaron (2014) by solving the model using the linearized version of the volatility dynamics to obtain quasi-closed form solutions for the linearization coefficients; for the inference of moments we use the original specification to ensure that the volatility of the model stays positive.

where the vol-of-vol factor q_t drives the volatility,[10]

$$\begin{aligned} \sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \phi_\sigma\sqrt{q_t}\omega_{\sigma,t+1} \\ q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\omega_{q,t+1} \\ x_{t+1} &= \rho x_t + \phi_x\sqrt{q_t}\eta_{x,t+1} \\ &\quad \eta_{x,t+1}, \omega_{\sigma,t+1}, \omega_{q,t+1} \sim i.i.d.\ N(0,1). \end{aligned} \tag{1.17}$$

The vol-of-vol factor q_t follows a square root process. This process specification has also been used, for example, in Tauchen (2011) or the seminal work on volatility of volatility in this model class by Bollerslev, Tauchen, and Zhou (2009). However, a square root process poses a new challenge to the model, as the process can become complex when q_t becomes negative. This problem is usually circumvented by assuming a reflecting boundary at zero to ensure positivity. (In fact, this approach has also been used for the stochastic volatility process in the original Bansal and Yaron (2004) study and many subsequent papers in the long-run risk literature.) However, for a simple computation of model solutions, the assumption of a non-truncated distribution for the log-linearization is commonly used. In Appendix 1.C we analyze in more detail how the square-root process specification and the issue of complexity affects the log-linearized solution. In particular we find that for the calibration in Bollerslev, Tauchen, and Zhou (2009) equilibrium model solutions are not real numbers but instead are complex numbers. For the parameters in Bollerslev, Xu, and Zhou (2015) the process is centered well above zero and the standard log-linearization technique yields a real solution. Therefore, we concentrate on this calibration in the main text.

The fifth study under consideration is the work on real and nominal bonds and the size of the martingale component in the stochastic discount factor by Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). They add inflation π_t with a stochastic growth rate $x_{\pi,t}$ to the standard model (1.13) and price nominal bonds[11]

$$\begin{aligned} \pi_{t+1} &= \mu_\pi + x_{\pi,t} + \phi_{\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_\pi\eta_{\pi,t+1} \\ x_{\pi,t+1} &= \mu_{x_\pi}(1-\rho_\pi) + \rho_\pi x_{\pi,t} + \rho_{\pi,x}x_t \\ &\quad + \phi_{x_\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{x_\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_{x_\pi}\eta_{\pi,t+1} \\ &\quad \eta_{\pi,t+1} \sim i.i.d.\ N(0,1). \end{aligned} \tag{1.18}$$

Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) assume that there are two stochastic volatility processes for consumption growth and the long-run risk component ($\sigma_{d,t} = \sigma_{c,t}$)

$$\sigma_{i,t+1}^2 = \bar{\sigma}_i^2(1-\nu_i) + \nu_i\sigma_{i,t}^2 + \phi_{\sigma_i}\omega_{i,t+1}, \qquad i \in \{c,d\},$$

[10] Drechsler and Yaron (2011) use a similar model where the volatility of x_t is driven by σ_t instead of q_t, see their 2007 working paper version. However, Bollerslev, Xu, and Zhou (2015) provide evidence for a better empirical match for their model specification. The estimation study of Bollerslev, Xu, and Zhou (2015) also models cross-correlations between the shocks of the state processes. For the analysis of the non-linear dynamics of the model we keep the model as parsimonious as possible and drop the cross-correlations.

[11] The model setup is the same as in the 2008 version of Bansal and Shaliastovich (2013). In the paper they write $\bar{\pi}_t$ for $x_{\pi,t}$.

and inflation, the stochastic growth rate of inflation and dividends have loadings on these two volatility channels.

The sixth and last study under consideration is the subsequent work on nominal and real bonds of Bansal and Shaliastovich (2013). The setup is very similar to Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) but they assume that $x_{\pi,t}$ enters the real stochastic growth rate of consumption x_t to model the non-neutral effect of expected inflation on future expected growth,

$$\begin{aligned} \pi_{t+1} &= \mu_\pi + x_{\pi,t} + \sigma_\pi \eta_{\pi,t+1} \\ x_{\pi,t+1} &= \rho_\pi x_{\pi,t} + \sigma_t^\pi e_{\pi,t+1} \\ x_{t+1} &= \rho x_t + \rho_{x\pi} x_{\pi,t} + \sigma_t^x e_{x,t+1} \\ &\eta_{\pi,t+1}, e_{\pi,t+1}, e_{x,t+1} \sim i.i.d.\ N(0,1). \end{aligned} \tag{1.19}$$

Also they assume that there is a separate AR(1) process for the volatility of the stochastic growth rate of inflation σ_t^π and the volatility of consumption growth is constant ($\sigma_{c,t} = \bar{\sigma}_c$):

$$\sigma_{i,t+1}^2 = \bar{\sigma}_i^2(1-\nu_i) + \nu_i \sigma_{i,t}^2 + \phi_{\sigma_i}\omega_{i,t+1}, \qquad i \in \{x, \pi\}.$$

As the focus of Bansal and Shaliastovich (2013) is on bond markets, they do not include a process for dividends.

Table 1.2 lists the parameter values of the six studies.[12] While the parameters in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a) are calibrated, Schorfheide, Song, and Yaron (2014), Bollerslev, Xu, and Zhou (2015) and Bansal and Shaliastovich (2013) estimate the model parameters to match annual financial market characteristics. In the first five models the investor has a monthly decision interval, while Bansal and Shaliastovich (2013) use quarterly intervals. This distinction explains, for example, the considerable difference in the level parameters. The main difference between the sets of parameter of the original Bansal and Yaron (2004) calibration and the new calibration of Bansal, Kiku, and Yaron (2012a) is that in the new calibration, the persistence of the volatility shock, ν_c, is higher and that shocks to dividends are correlated with short-run shocks to consumption growth ($\phi_{d,c} = 2.6$ in the new calibration compared to $\phi_{d,c} = 0$ in the original calibration). These changes increase the influence of the volatility channel compared to the long-run risks channel of the model. The adjustment is needed to get rid of some implications of the original calibration that are inconsistent with the data. In particular, as, for example, Zhou and Zhu (2015) or Beeler and Campbell (2012) point out for the original 2004 calibration, the log price-dividend ratio has predictive power for future consumption growth, while this relationship is not present in the data. By increasing the influence of the volatility channel, this predictability vanishes.

[12]For the model of Bollerslev, Xu, and Zhou (2015) we use the parameters estimates in the study for ρ, ν and ρ_q. As they do not report values for the remaining parameters, we use the calibration as reported in the 2007 working paper version of Drechsler and Yaron (2011).

Table 1.2: Model Parameters

		BY (2004)	BKY (2012)	SSY (2014)	BS (2013)	KLVV (2010)	BXZ (2015)
Preferences	γ	10	10	10.84	20.90	8	10
	ψ	1.5	1.5	1.7	1.81	1.5	1.5
	δ	0.998	0.9989	0.9996	0.994	0.9987	0.999
Cons.	μ_c	0.0015	0.0015	0.0016	0.0049	0.0016	0.0015
	ϕ_c	1	1	1	1	1	0.00546
	ρ	0.979	0.975	0.993	0.81	0.991	0.988
	ϕ_x	0.044	0.038	1	1	1	3.12e–4
	$\rho_{x\pi}$	–	–	–	–0.047	0	–
Volatility	ν_c	0.987	0.999	0.956	0	0.85	0.64
	ν_x	–	–	0.99	0.994	0.996	–
	ν_d	–	–	0.94	–	-	–
	ν_π	–	–	–	0.979	-	–
	ϕ_{σ_c}	2.3e–6	2.8e–6	8.8e–6	0	1.15e–6	1
	ϕ_{σ_x}	–	–	6.0e–9	1.85e–7	4.19e–9	–
	ϕ_{σ_d}	–	–	2.3e–4	–	-	–
	ϕ_{σ_π}	–	–	–	1.81e–7	-	–
	$\bar{\sigma}_c$	0.0078	0.0072	0.005	4.6e–3	0.004	1
	$\bar{\sigma}_x$	–	–	2.0e–4	1.09e–3	1.60e–5	–
	$\bar{\sigma}_d$	–	–	0.0273	–	-	–
	$\bar{\sigma}_\pi$	–	–	–	1.11e–3	-	–
Dividends	μ_d	0.0015	0.0015	0.001	–	0.0015	0.0015
	Φ	3.0	2.5	3.2	–	1.5	3.0
	ϕ_d	4.5	5.96	1	–	6	0.0246
	$\phi_{d,c}$	0	2.6	1.17	–	0.6	0
Inflation	μ_π	–	–	–	0.0090	0	–
	μ_{x_π}	–	–	–	0	0.0032	–
	σ_π	–	–	–	0.0055	0.0035	–
	σ_{x_π}	–	–	–	0	4e–6	–
	$\phi_{\pi,c}$	–	–	–	0	0	–
	$\phi_{\pi,x}$	–	–	–	0	-2	–
	$\phi_{x_\pi,c}$	–	–	–	0	0	–
	$\phi_{x_\pi,x}$	–	–	–	0	-1	–
	ρ_π	–	–	–	0.988	0.83	–
	$\rho_{\pi,x}$	–	–	–	0	-0.35	–
Vol–of–Vol	μ_q	–	–	–	–		0.211
	ϕ_q	–	–	–	–		0.632
	ρ_q	–	–	–	–		0.46

Parameter values as reported in the studies of Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012a), Schorfheide, Song, and Yaron (2014), Bansal and Shaliastovich (2013), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), and Bollerslev, Xu, and Zhou (2015).

Schorfheide, Song, and Yaron (2014) provide further evidence for a highly persistent stochastic growth rate $\rho = 0.993$ with a 90% confidence interval of $\{0.989, 0.994\}$. In line with the calibrated values in Bansal, Kiku, and Yaron (2012a), they also find a highly persistent volatility process for the long-run risk component, while the estimates for consumption and dividend volatility are slightly smaller.

1.4.2 Moments and Errors

Table 1.3 reports annualized summary statistics and numerical errors for the five models that include a dividend process. The reported financial statistics are the mean and standard deviation of the price-dividend ratio, the averages of the market excess return and the risk-free return, and the volatilities of the excess return and the risk-free rate.[13] The table reports these statistics for both the solution of the log-linearization approach and the projection approach; in addition, it states the relative errors induced by the linearization.

We observe that the log-linearization does a reasonably good job for the parameters in Bansal and Yaron (2004) with a maximal error of 2.93% for the equity premium. For the parameter set of Bansal, Kiku, and Yaron (2012a) the results are considerably worse. The log-linearization overstates the equity premium by more than 100 basis points; and it predicts a volatility of the log price-dividend ratio of 0.2910 instead of 0.2389. These values correspond to relative errors of about 22%. Simply put, the log-linearization falsely produces a large equity premium and volatile log price-dividend ratio even though the true model solution is significantly smaller.

For the model of Schorfheide, Song, and Yaron (2014) approximation errors become even larger. For the equity premium and the risk-free rate the errors exceed 50% and also the errors in the other four key statistics exceed 10%.[14] In Section 1.4.4 below, we carve out the source for these large numerical errors. It is the interplay of the highly persistent state processes that introduces substantial non-linearities to the model solutions; as a result, even a slight increase in the persistence parameter of the long-run risk channel can dramatically increase the approximation errors of the log-linearized solution. Schorfheide, Song, and Yaron (2014) estimate a persistence of $\rho = 0.993$ compared to $\rho = 0.975$ in the calibration of Bansal, Kiku, and Yaron (2012a)

[13] To compute the annualized moments, we simulate 1,000,000 years of artificial data. Beeler and Campbell (2012) provide a detailed description of how to compute the annual moments from the monthly observations. A significant issue in the model is that the variance process σ_t^2 can, in fact, become negative. To overcome this problem, Bansal and Yaron (2004) replace all negative realizations with very small but positive values. We proceed in the same way for both methods to achieve consistent results. For the approximation interval of the projection methods we choose the interval to be slightly larger than the maximum observation range of the long simulations. As in the previous section, we increase the polynomial degree until the coefficients of the highest-order polynomial are close to zero. We double-check the accuracy of the solution by increasing the approximation interval until the solutions do not change.

[14] Note that the results are very sensitive to changes in the model parameters. Here we show model outcomes for the median estimates of Schorfheide, Song, and Yaron (2014), while in the original study they draw parameter values from the estimated distributions of the model parameters and report the median for a large number of draws. For example, for the 5% quantile estimates the model yields an equity premium of 2.4% with a risk-free rate of 2.3%. This explains why the values reported here differ from the values shown in Table 4 of the study of Schorfheide, Song, and Yaron (2014).

Table 1.3: Annualized Moments and Errors

	$E(p_t - d_t)$	$\sigma(p_t - d_t)$	$E\left(r_t^m - r_t^f\right)$	$E\left(r_t^f\right)$	$\sigma(r_t^m)$	$\sigma\left(r_t^f\right)$
Bansal and Yaron (2004)						
Log-Lin	3.1749	0.2012	4.61	1.46	17.05	1.31
Projection	3.2056	0.1990	4.48	1.46	16.97	1.31
Error	0.96 %	1.12%	2.93%	0.08%	0.50%	0.05%
Bansal, Kiku, and Yaron (2012a)						
Log-Lin	3.0473	0.2910	5.73	0.99	21.27	1.28
Projection	3.2413	0.2389	4.69	1.10	21.00	1.27
Error	5.98%	21.81%	22.26%	10.21%	1.28%	1.45%
Schorfheide, Song, and Yaron (2014)						
Log-Lin	1.9394	0.3331	18.00	-1.80	20.43	1.56
Projection	2.3497	0.2892	12.00	-1.17	18.43	1.39
Error	17.46%	15.18 %	50.02%	54.54%	10.84%	12.29%
Bollerslev, Xu, and Zhou (2015)						
Log-Lin	2.7479	0.2485	7.27	1.16	16.28	1.36
Projection	2.8225	0.2399	6.78	1.17	15.91	1.35
Error	2.64%	3.58%	7.26%	0.72%	2.35%	0.40%
Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010)						
Log-Lin	3.1102	0.1782	4.85	1.64	11.53	1.17
Projection	3.3468	0.1465	3.56	1.32	10.58	1.14
Error	7.07%	21.66%	36.29%	19.43%	9.07%	2.52%

The table shows the mean and the standard deviation of the annualized log price-dividend ratio, the annualized market over the risk-free return and the risk-free return. Results obtained by the log-linearization and the projection method as well as the relative error of the log-linearization are shown for the models of Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012a), Schorfheide, Song, and Yaron (2014), Bollerslev, Xu, and Zhou (2015) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). All returns are shown in percent, so a value of 1.5 is a 1.5% annualized figure.

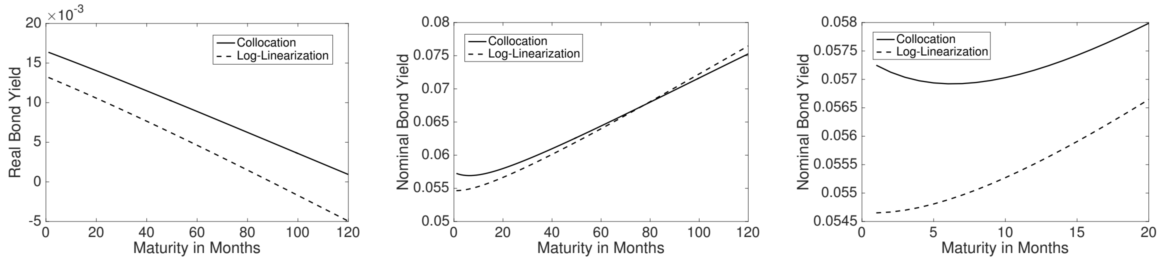
which explains the large approximations errors. Hence, using the log-linearized solution to estimate models featuring highly persistent state processes can potentially introduce a large bias to the implied model moments and so, in turn, biases the estimation results for the model parameters.

This finding is in line with the results for the model of Bollerslev, Xu, and Zhou (2015). The model only features a highly persistent long-run risk process $\rho = 0.988$ while the persistences of the stochastic volatility and vol-of-vol factors are rather low ($\nu = 0.64$ and $\rho_q = 0.46$). Consequently the approximation errors are rather small with a maximum error of 7.26% for the equity premium. This result is not surprising as the authors mention in their estimation that the stochastic volatility and the vol-of-vol factors only influence the variance premium and have a negligible influence on the price and return dynamics. Concordantly, we obtain almost the same results when setting the volatility of the two factors to zero ($\phi_\sigma = \phi_q = 0$).

For the study of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) we also find large errors with a maximum error in the equity premium of 36.29%. An overestimation of the premium of more than 100 basis points. Their calibration features a highly persistent long-run risk process $\rho = 0.991$ and highly persistent stochastic volatility of long-run risk $\nu_x = 0.996$ that introduce the large non-linearities to the model. Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) not only analyze equity markets but also price real and nominal bonds to analyze the martingale component in the stochastic discount factor. In Figure 1.4 we show the real and nominal yield curve for their model.

Figure 1.4: Real and Nominal Yield Curve in the Model of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010)

(a) Real YC (b) Nominal YC (c) Nominal YC (1-20)



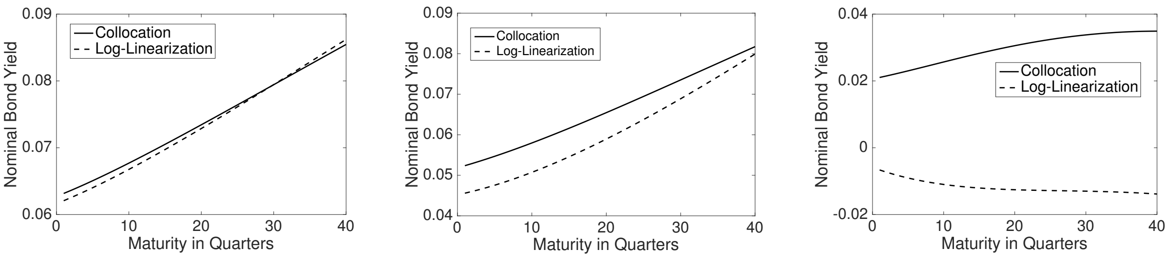
The graph shows the yield curves for real and nominal bonds in the model of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). Panel (c) shows the yield curve for 1-20 months bonds.

We find that the differences between the yield curve obtained by linearizing the model and solving it accurately using the projection approach are small in absolute values. However, the nominal yield curve from the linearized model differs in its shape. While the true nominal yield curve is downwards sloping in the short run and upwards sloping in the long run, this pattern does not occur when using log-linearization. So linearizing the model potentially affects the shape of the real curve.

The work of Bansal and Shaliastovich (2013) provides further insights to this finding. In Figure 1.5 we show the nominal yield curve in their model.

Figure 1.5: Nominal Yield Curve in the Model of Bansal and Shaliastovich (2013)

(a) $\rho = 0.81$ (b) $\rho = 0.9$ (c) $\rho = 0.975$



The graph shows the yield curve for nominal bonds in the model of Bansal and Shaliastovich (2013).

Panel (a) shows the yield curve for the parameters in the original study. We observe that the difference between the log-linearized solution and the projection solution is negligible with very small errors and also the shape of the yield curve is correct. As Bansal and Shaliastovich (2013) use bond data to estimate the model, they find a very low persistence in the long-run risk component with $\rho = 0.81$. This comparably low amount of persistence makes it difficult to match key moments for equity markets. For example the annualized equity premium for their parameter estimates is only 1.69%.[15] Therefore we increase ρ in panels (b) and (c) to 0.9 and 0.975 correspondingly to increase the premium paid for long-run consumption risk.[16] We find that the errors in the yield curve grow significantly as ρ approaches the value 1. In fact, for $\rho = 0.975$ the log-linearization predicts a downward sloping nominal yield curve (dashed line) even though the model actually produces an upward sloping curve (solid line). Hence, relying on the log-linearization to solve the model can lead to false conclusions not only about the magnitude of bond yields but even about the shape of the yield curve.

In sum, we observe that while the log-linearization approach produces satisfactory solutions for an analysis of the models in Bansal and Yaron (2004) and Bollerslev, Xu, and Zhou (2015), the method performs rather poorly for the models in Bansal, Kiku, and Yaron (2012a), Schorfheide, Song, and Yaron (2014), and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). For these latter models, the poor approximations have a strong effect on the model predictions for key financial statistics. Our observations motivate the next step in our analysis. We want to understand which model characteristics affect the performance of the log-linearization approach; simply put, when can we trust the results of such an approach and when can we

[15]The published version of Bansal and Shaliastovich (2013) does not provide a process for dividend growth. For the purposes of comparison, we consider the specification that appears in the 2007 working paper of their paper. The process for Δd_{t+1} is the same as in Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) (see equation 1.18). As the 2007 working paper assumes a monthly decision interval and the published version from 2013 has a quarterly interval, we adjust the volatility of dividends ϕ_d to match the volatility of dividend growth in the data of approximately 11% annualized.

[16]For $\rho = 0.9$ we obtain an equity premium of 4.48% and for $\rho = 0.97$ a premium of 10.57%.

not? And related to this question, we also want to understand which properties of the exact solution lead to a poor performance of a linear method; that is, what exactly goes wrong with the linearized solution?

1.4.3 The Interplay of the State Processes

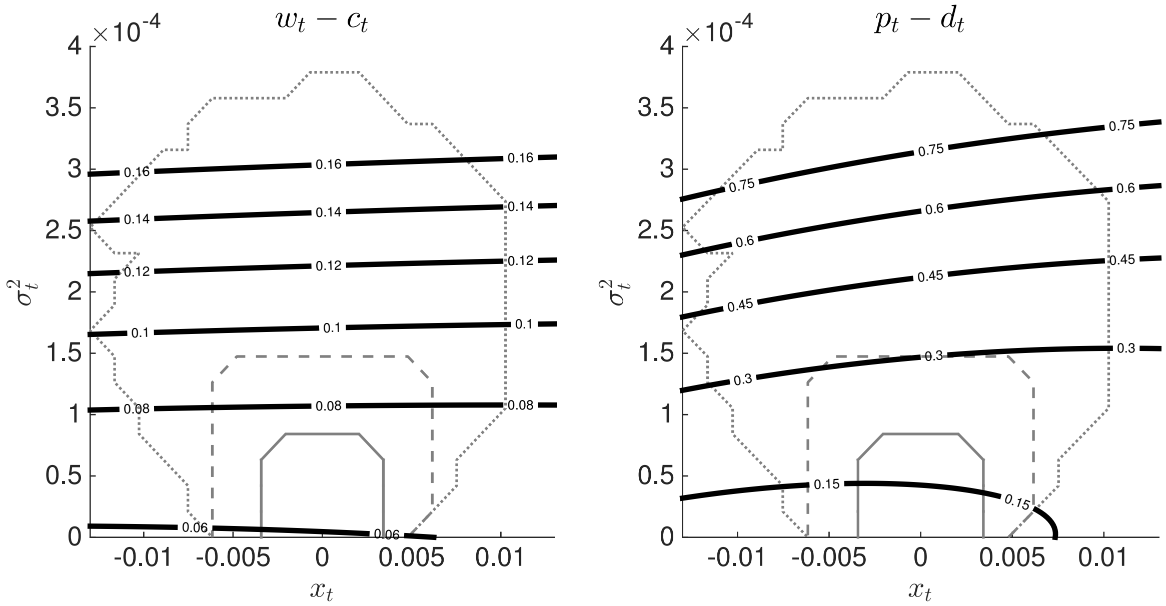
The log-linearization approach assumes that, on the state space of the model, the first derivatives of the solution are approximately constant and the second derivatives are approximately zero. We now show numerically that this assumption fails to hold for models with more than one highly persistent state process. We demonstrate that for solutions of such models the second derivatives can be very large, the interplay of state process leads to highly non-linear solutions; and so higher-order effects matter for the predictions of such models. The sizable deviations from linearity in the models' solutions is the cause for the failure of the log-linearization approach.

For the purpose of making these points, we concentrate on the two fundamental factors of long-run risk and stochastic volatility. We use the calibration of Bansal, Kiku, and Yaron (2012a) (see equations (1.13)-(1.14)). Figure 1.6 shows isolines for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x and σ^2 (black solid lines). For example along a line marked with '0.1', the absolute error of the log-linearization is 0.1. The figure also shows the regions into which 50%, 90% and 100% of the observations fall. These regions show the subsets of the state space that the model actually visits and in which regions it "spends most of its time" during long simulations. Corresponding errors for the first derivatives with respect to the state variables are shown in Figure 1.7 and for the second derivative in Figure 1.8.

We find that the errors in the log wealth-consumption are rather small with maximum values of about 0.16 within the observation range. For the log price-dividend ratio, the errors are also small in the area close to the long-run mean of the processes, but they increase significantly with σ^2 and reach values of up to 0.3 in the 90% observation range, see Figure 1.6. Put differently, the price dividend ratio obtained by the log-linearization is off by a factor of $e^{0.3} \approx 1.35$ for almost 10% of the time and can be off by a factor larger than 2 for extreme values reached in the simulations.

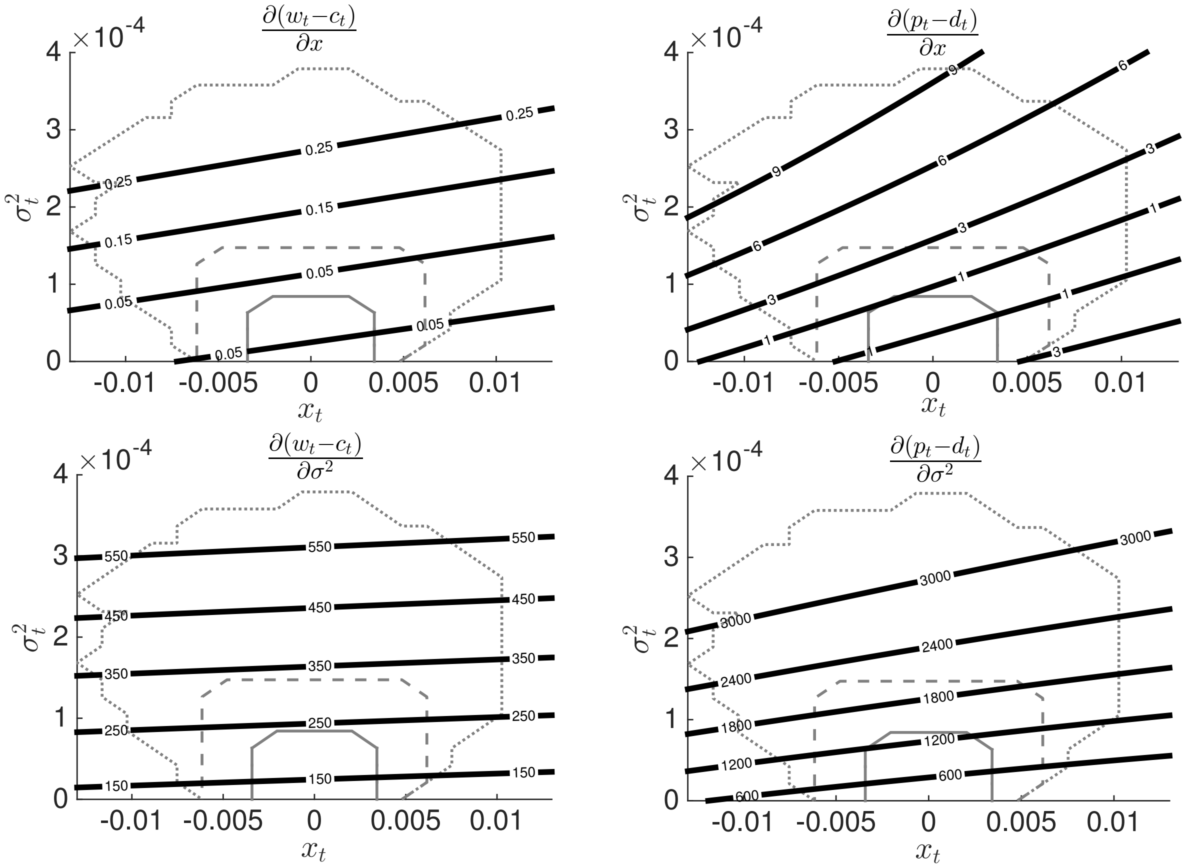
The errors in the first derivatives show similar patterns. Again the errors in the derivatives of the price-dividend ratio are significantly larger than the errors in the derivatives of the wealth-consumption ratio and the errors increase monotonically with σ^2 for the BKY (2012) calibration. We observe in Figure 1.7 that the errors in the derivatives with respect to σ^2 are especially large, with errors up to 3000 for the price-dividend ratio. As mentioned above, the main purpose of the BKY (2012) calibration is to amplify the role of the stochastic volatility channel by increasing its persistence. But as demonstrated in the figures, this effect introduces large non-linearities to the model that cannot be captured by the log-linearization and hence causes large approximation errors. Figure 1.8 shows that the second derivatives in the model are substantially different from 0 (which is the value assumed by the log-linearization) and they are especially large (more than 10^5!) for the second derivative with respect to σ^2 which is another reason for the large approximation errors reported in Table 1.3.

Figure 1.6: Approximation Errors in the log Wealth-Consumption and log Price-Dividend Ratio of the Log-Linearization



The graph shows isolines for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x and σ^2 (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from 10^6 simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2.

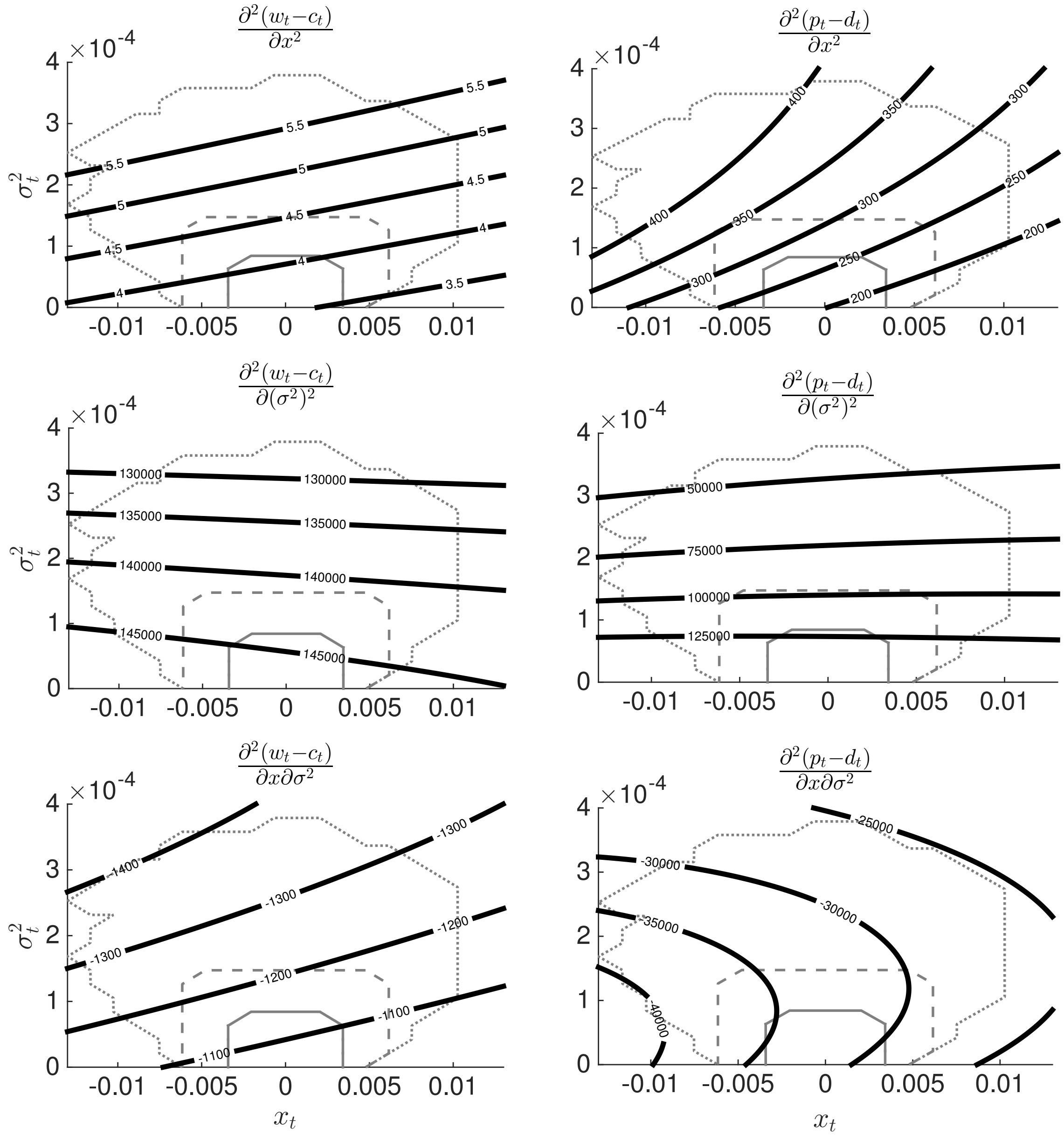
Figure 1.7: Approximation Errors in the First Derivatives of the log Wealth-Consumption and log Price-Dividend Ratio of the Log-Linearization



The graph shows isolines for the absolute errors in the first derivative of the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) with respect to the states x and σ^2 of the log-linearization (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from 10^6 simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2.

In general, the figures show that the stochastic volatility channel highly influences the non-linear aspects of the model. But is it only the stochastic volatility that matters? Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012) analyze the accuracy of several solution methods in a neoclassical growth model with Epstein-Zin preferences and stochastic volatility. They report that higher-order approximations are needed to capture the non-linearities of the model. Bansal, Kiku, and Yaron (2012b) report approximation errors for the long-run risks model in their estimation study by comparing the results of the log-linearization to the results obtained by the discretization method of Tauchen and Hussey (1991) (see Table A.1 of their paper). Unfortunately, for their exercise, they use a simplified version of their model that only features long-run risks (and no stochastic volatility). They find rather small approximation errors. But in the long-run risk model, there are two sources of non-linearities: the stochastic volatility channel and the long-run risk channel. Hence when solving the model, it is essential to understand whether and how the interplay of the two components drives the non-linearities.

Figure 1.8: Approximation Errors in the Second Derivatives of the log Wealth-Consumption and Price-Dividend Ratio



The graph shows isolines for the absolute errors in the second derivative of the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) with respect to the states x and σ^2 of the log-linearization (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from 10^6 simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2.

To obtain such an understanding, we analyze the approximation errors implied by the log-linearization for each of the two state variables of the model separately. In particular we first fix the stochastic volatility to its long-run mean, $\sigma_t = \bar{\sigma}^2 \ \forall\ t$, and secondly we solve the model without long-run risk, $x_t = 0 \ \forall\ t$.

Table 1.4: Approximation Errors for Each State of the Long-Run Risks Model Separately

	$E\,(w_t - c_t)$	$\sigma\,(w_t - c_t)$	$E\,(p_t - d_t)$	$\sigma\,(p_t - d_t)$
State: x_t	0.003%	0.024%	0.084%	0.21%
State: σ_t^2	0.14%	4.49%	2.62%	7.05%
Both States:	1.05%	12.25%	3.15%	26.90%

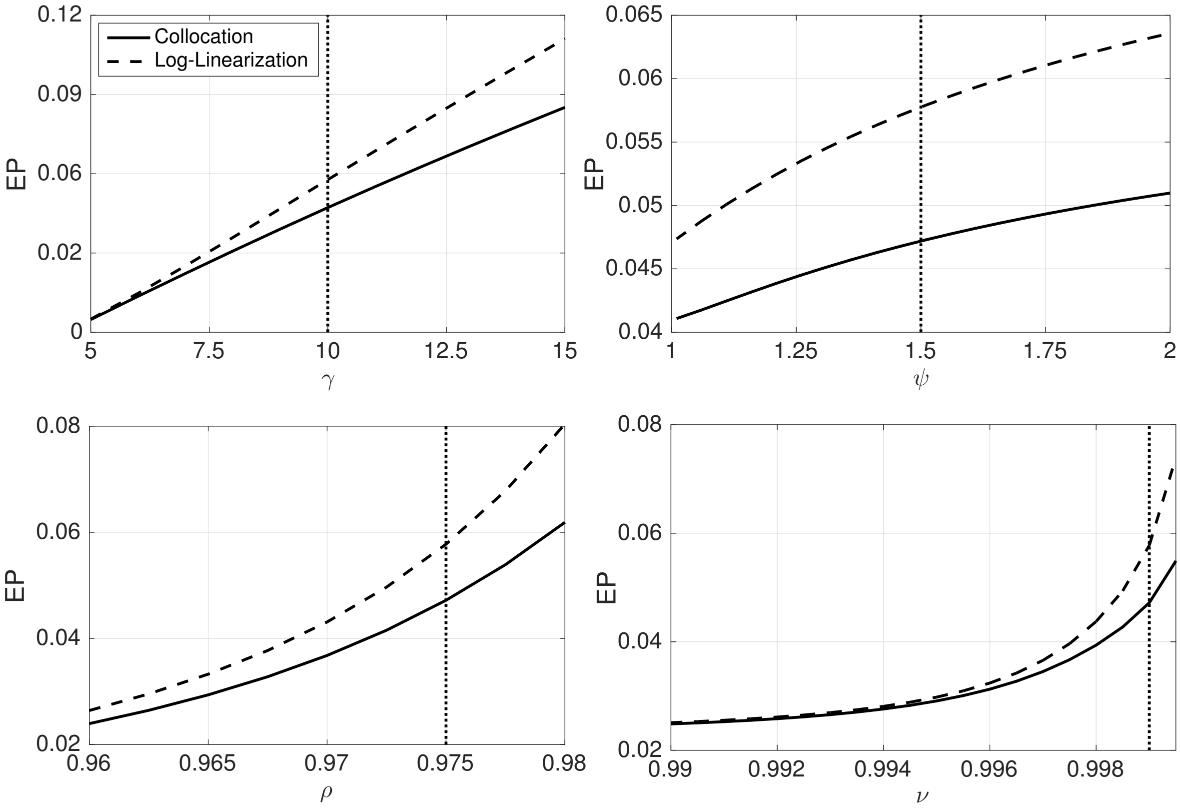
The table shows approximation errors in the unconditional mean and standard deviation of the log wealth-consumption and log price-dividend ratio induced by the log-linearization in the long-run risk model for each of the two state variables x_t and σ_t separately. For the case with only x_t, the state σ_t is simply set constant at its long-run mean $\bar{\sigma}^2$ (or equivalently $\nu = \sigma_w = 0$). For the case with only σ_t, x_t is set to 0 (or equivalently $\rho = \phi_x = \Phi = 0$). The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2.

Table 1.4 shows the corresponding errors in the unconditional mean and standard deviation of the log wealth-consumption and log price-dividend ratio for the two cases. We find that, in line with the test results from Bansal, Kiku, and Yaron (2012b), for the one-dimensional model with only long-run risks the approximation errors are very small with a maximum error of 0.21%). For the second case, without long-run risks and only stochastic volatility, the errors are slightly larger but still remain below 7.1%. However, for the full model with long-run risk and stochastic volatility approximation errors increase dramatically with a maximum error of 26.9% for the volatility of the log price-dividend ratio. This finding suggests that neither the stochastic volatility alone nor the long-run risks component alone introduces the non-linearities in the model; instead it is the simultaneous presence and interplay of the two features which makes the model so difficult to solve.

1.4.4 Sensitivity of the Approximation Errors

As the previous results have shown, the non-linearities of the long-run risk model are highly dependent on its parameters. Therefore, in Figures 1.9 and 1.10 we analyze the approximation errors implied by the log-linearization with regard to changes in the parameters. In particular, we consider those parameters that are the main driving forces of the model, namely the risk aversion, γ, the intertemporal elasticity of substitution, ψ, the serial correlation in the long-run risk channel, ρ, and the stochastic volatility channel, ν.

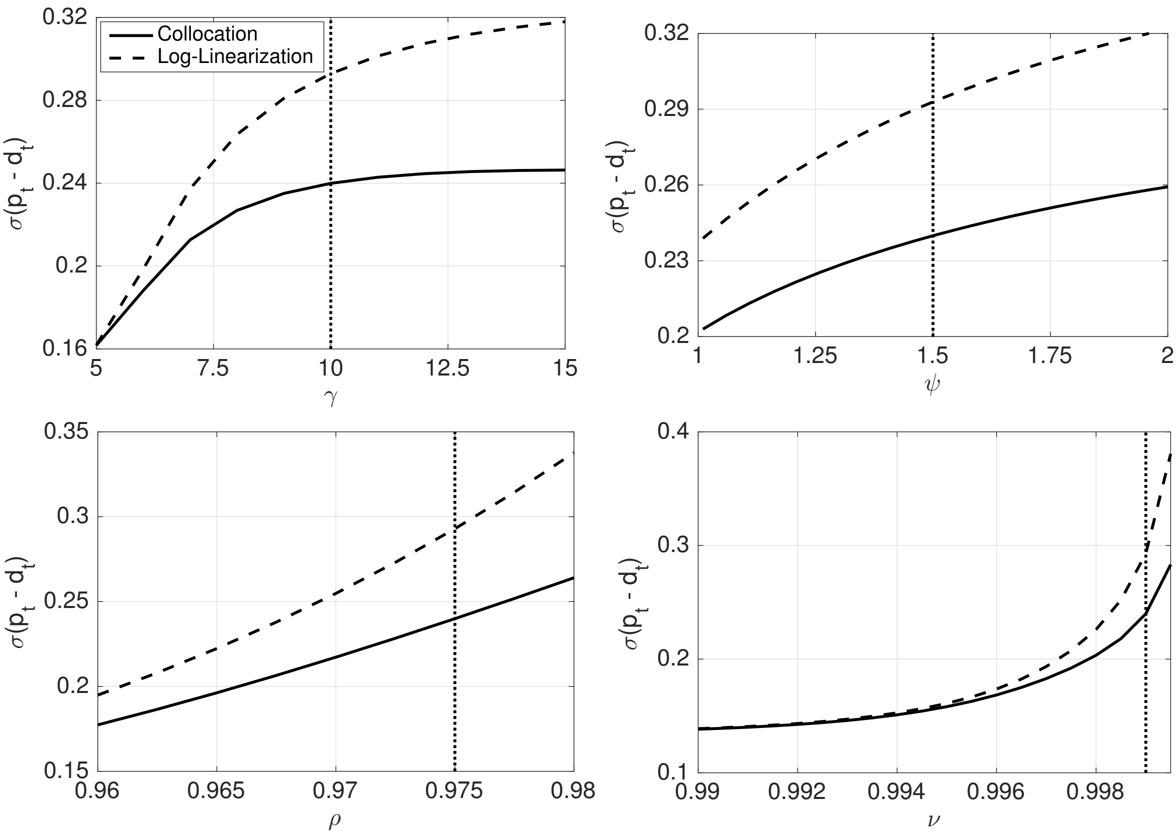
Figure 1.9: Sensitivity of the Approximation Errors for the Equity Premium in the Long-Run Risks Model



The figure shows the equity premium obtained by the log-linearization (dashed line) as well as the premium obtained by the collocation projection (solid line) as a function of the model parameters γ, ψ, ρ and ν, respectively, assuming that the other parameters are kept constant. The results are computed for the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2, and in each panel, the dotted vertical line denotes the estimate used in original calibration.

We find that, for this particular calibration, for a risk aversion of approximately 5, the log-linearized solution basically coincides with the solution from the projection approach, which suggests that a linear solution gives a reasonable approximation to the model. However, for this calibration also the implied model moments collapse with an equity premium below 1% and a sharp decrease in the volatility of the log price dividend ratio. When increasing the risk aversion the errors in the equity premium and the volatility of the log price-dividend ratio increase significantly, with a large overestimate of both quantities. Furthermore, in line with the previous results, the accuracy depends highly on the persistence of the processes for both the long-run risk and the stochastic volatility. We observe that even very small changes can dramatically increase approximation errors. For example, in the original calibration with a persistence in the long-run risk of $\rho = 0.975$ the overestimation of the equity premium is about 100 basis points (see Table 1.3). By slightly increasing ρ to 0.98, however, the difference doubles with an overestimation of 200 basis points. This very strong dependence on the persistence parameters also explains the large approximation errors in the estimation study of

Figure 1.10: Sensitivity of the Approximation Errors for the Volatility of Price-Dividend Ratio in the Long-Run Risks Model



The figure shows the volatility of the log price-dividend ratio obtained by the log-linearization (dashed line) as well as the volatility obtained by the collocation projection (solid line) as a function of the model parameters γ, ψ, ρ and ν, respectively, assuming that the other parameters are kept constant. The results are computed for the calibration of Bansal, Kiku, and Yaron (2012a), see Table 1.2, and in each panel, the dotted vertical line denotes the estimate used in original calibration.

Schorfheide, Song, and Yaron (2014) (see Table 1.3) that finds a serial correlation in the long-run risk channel of $\rho = 0.993$. For the persistence in the conditional variance, ν, even a change of 0.0005, from 0.999 to 0.9995, increases the overestimation to 200 basis points. The figures also show that lowering the persistence parameters significantly decreases approximation errors. For example for $\nu = 0.99$ the approximation error becomes close to zero. However, for this calibration also the implied model moments collapse. Therefore it is especially important to pay attention to accurately solving the model as small changes to the parameters can have large impacts on the higher-order dynamics and hence introduce large approximation errors when using log-linear approximations; thus further applications of this class of models, require robust and accurate solution methods like the projection method presented in this paper.

1.5 Conclusion

We have investigated the existence of solutions for long-run risk models and the accuracy of the Campbell-Shiller log-linear approximation to those solutions. For existence, we have provided a relative existence result – if the model has a solution for CRRA preferences, then it has a solution for investors with a preference for an early resolution of uncertainty. Existence can be proven for the Bansal and Yaron (2004) model with CRRA preferences, so existence for early resolution follows.

To evaluate the quality of the log-linear solutions, we consider six recent models in the long-run risk literature: the original Bansal and Yaron (2004) model and the new calibration of Bansal, Kiku, and Yaron (2012a), the estimation of Schorfheide, Song, and Yaron (2014), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015) and the work on real and nominal bonds of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013). We find for very persistent underlying processes the approximation errors in log-linearization can be large and economically significant. For example, in the most recent calibration of the Bansal-Yaron long-run risk model (see Bansal, Kiku, and Yaron (2012a)), the approximation errors in the volatility of the log-price dividend ratio and the equity premium exceed 22% and become as large as 50% in the estimation study of Schorfheide, Song, and Yaron (2014). Models with lower persistence, such as the original Bansal and Yaron (2004) model or Bollerslev, Xu, and Zhou (2015), have much smaller approximation errors. The results for nominal bonds as in Bansal and Shaliastovich (2013) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) are particularly interesting – for the high level of persistence necessary to explain the equity premium, the log-linear approximation can actually produce an downward sloping yield curve, when the true yield curve is upward sloping.

Given the importance of long-run risk model in asset pricing, our results suggest that more sophisticated solution methods, such as projection methods, should be used when it comes to asset pricing models with highly persistent state processes.

1.A Proofs for Section 1.3

Proof of Lemma 1. If $f \leq g^*$, then $Tf \leq Tg^* \leq Ug^* = g^*$. So, T maps $(0, g^*]$ into itself. By assumption, T is monotone. And so Tarski's (1955) Fixed-Point Theorem implies that T has a fixed point in the complete lattice $(0, g^*]$. □

Proof of Lemma 2. If $\theta < 1$, then by Jensen's inequality if $E(f|s)$ is finite, then $E(f^\theta|s)$ is finite. If $\theta > 1$, then again by Jensen's inequality if $E(f^\theta|s)$ is finite then $E(f|s)$ is finite. In both cases, T and U preserve $\mathcal{V}$.

(A) Let $0 \neq \theta \leq 1$. If $\theta > 0$, then x^θ is increasing in $x \in \mathbb{R}_{++}$. The monotonicity of the expected value operator implies for $f \leq g$ that

$$\begin{aligned} E(f^\theta|s) &\leq E(g^\theta|s), \\ \left[E(f^\theta|s)\right]^{1/\theta} &\leq \left[E(g^\theta|s)\right]^{1/\theta}, \end{aligned}$$

and thus $Tf \leq Tg$. If $\theta < 0$, then x^θ is decreasing in $x \in \mathbb{R}_{++}$. Now we obtain

$$\begin{aligned} E(f^\theta|s) &\geq E(g^\theta|s), \\ \left[E(f^\theta|s)\right]^{1/\theta} &\leq \left[E(g^\theta|s)\right]^{1/\theta}, \end{aligned}$$

and thus $Tf \leq Tg$ in this case as well. Trivially, $1/\theta \geq 1$ for $0 < \theta \leq 1$ and $1/\theta < 0$ for $\theta < 0$. In both cases, $x^{1/\theta}$ is convex for $x \in \mathbb{R}_{++}$. Therefore, by Jensen's inequality

$$E(X|s) = E((X^\theta)^{1/\theta}|s) \geq \left[E(X^\theta|s)\right]^{1/\theta}.$$

Thus, for $0 \neq \theta \leq 1$ it holds that $Tg \leq Ug$.

(B) Let $\theta > 1$. The monotonicity of the expected value operator implies for $f \leq g$ that $Uf \leq Ug$. Trivially, $1/\theta < 1$ and so $x^{1/\theta}$ is now concave for $x \in \mathbb{R}_{++}$. Therefore, by Jensen's inequality

$$E(X|s) = E((X^\theta)^{1/\theta}|s) \leq \left[E(X^\theta|s)\right]^{1/\theta}.$$

Hence, any pair $f, g \in \mathcal{V}$ with $f \leq g$ satisfies $Uf \leq Ug \leq Tg$.

□

Proof of Theorem 1.

(A) The asset pricing model characterized by equations (1.1)–(1.6) has a solution for CRRA utility if and only if equation (1.8) has a solution $\hat{V}_t = \hat{V}_{t+1} = \hat{V}^*$ for $\theta = 1$. Lemma 3 implies that the equation also must have a solution for $0 \neq \theta < 1$, that is, for Epstein-Zin utility. Since λ and thus ψ are fixed, the conditions on γ follow.

(B) The asset pricing model characterized by equations (1.1)–(1.6) has a solution for Epstein-Zin utility if and only if equation (1.8) has a solution $\hat{V}_t = \hat{V}_{t+1} = \hat{V}^*$ for $\theta \neq 1$. If $\theta > 1$, then Lemma 3 implies that the equation also must have a solution for $\theta = 1$, that is, for CRRA utility. Since λ and thus ψ are fixed, the conditions on γ follow.

□

Sketch of Proof of Theorem 2. The following lines are closely related to the work in de Groot (2015). In the online appendix de Groot (2015) shows how to derives closed-form solutions for the price-dividend ratio of the market portfolio in the model of Bansal and Yaron (2004) with CRRA preferences. He also presents a formal convergence theorem for the model without the long-run risk factor. Unfortunately he doesn't provide a convergence theorem for the pricing of the consumption claim for the full long-run risk model, which is needed for the existence of solutions for recursive preferences analyzed in this study. The following lines fill this gap.

Following Section B.4.2 in de Groot (2015) the solution of the wealth-consumption ratio $Z_{w,t} = \exp(z_{w,t})$ is of the form[17]

$$Z_{w,t} = \sum_{i=1}^{\infty} \delta^i \exp\left(A_{1,i} + A_{2,i}\bar{\sigma}^2 + A_{3,i}x_t + A_{4,i}\phi_x^2 + A_{5,i}(\bar{\sigma_t}^2 - \bar{\sigma}^2) + A_{6,i}\phi_\sigma^2\right).$$

The coefficients for the wealth-consumption ratio can be obtained in the same way as conducted by de Groot (2015) for the price-dividend ratio. To save space we do not state the full coefficients here. Please contact the authors for the detailed derivation of the coefficients. Define

$$Z_w^i \equiv \delta^i \exp\left(A_{1,i} + A_{2,i}\bar{\sigma}^2 + A_{3,i}x_t + A_{4,i}\phi_x^2 + A_{5,i}(\bar{\sigma_t}^2 - \bar{\sigma}^2) + A_{6,i}\phi_\sigma^2\right).$$

$Z_{w,t}$ is finite if

$$\lim_{i\to\infty} \left| \frac{Z_w^{i+1}}{Z_w^i} \right| < 1. \tag{1.20}$$

[17]Note that the notation is slightly different from the specification in de Groot (2015). While de Groot (2015) summarizes the constant terms $A_{2,i}$ and $A_{4,i}$ in one term called C_i^{BY} in the paper, we separate the two terms as the first term captures the influence of the short term shock to consumption growth, while the second term captures the influence of shock to the long-run growth rate.

Inserting the solutions for the coefficients we obtain

$$
\begin{aligned}
\lim_{i\to\infty} A_{1,i+1} - A_{1,i} &= \mu_c\left(1-\frac{1}{\psi}\right) \equiv B \\
\lim_{i\to\infty} A_{2,i+1} - A_{2,i} &= 0.5\left(1-\frac{1}{\psi}\right)^2 \equiv B_c \\
\lim_{i\to\infty} A_{3,i+1} - A_{3,i} &= 0 \\
\lim_{i\to\infty} A_{4,i+1} - A_{4,i} &= 0.5\left(\frac{1-\frac{1}{\psi}}{1-\rho}\right)^2 \bar{\sigma}^2 \equiv B_x \\
\lim_{i\to\infty} A_{5,i+1} - A_{5,i} &= 0 \\
\lim_{i\to\infty} A_{6,i+1} - A_{6,i} &= \frac{1}{8}\left(\frac{\left(1-\frac{1}{\psi}\right)^2\phi_x^2}{(1-\rho)(1-\nu)}\right)^2 + \frac{1}{4}\left(\frac{\left(1-\frac{1}{\psi}\right)^2\phi_x}{(1-\rho)(1-\nu)}\right)^2 \\
&+ \frac{1}{8}\left(\frac{\left(1-\frac{1}{\psi}\right)^2}{(1-\nu)}\right)^2 \equiv B_\sigma
\end{aligned}
$$

Hence there exists a solution for the wealth-consumption ratio $Z_{w,t}$ if and only if

$$\delta \exp\left(B + B_c\bar{\sigma}^2 + B_x\phi_x^2 + B_\sigma\phi_\sigma^2\right) < 1. \tag{1.21}$$

□

1.B Accuracy of the Projection Method

Table 1.5 demonstrates the accuracy of the projection approach. Therefore we use the long-run risk model of Bansal and Yaron (2004) with constant volatility where there exist closed form solutions for the case of CRRA preferences (Equations (1.7) with $\sigma_{c,t} = \sigma_{x,t} = \bar{\sigma}$ and $\eta_{c,t+1}$, $\eta_{x,t+1}$ i.i.d. normal.). In the case of recursive preferences we determine the accurate solution using the projection approach with a very large degree and state space. We use a state space of $n_\sigma = 50$ standard deviations around the unconditional mean of x_t and increase the polynomial approximation degree n until the highest order coefficient is close to zero. We double check the solution by using the discretization method of Tauchen and Hussey (1991) with a very large number of discretization nodes.

We find that for the calibration with $\rho = 0.95$ already a first order approximation with an approximation interval of $n_\sigma = 1$ standard deviation around the unconditional mean of x_t provides a very accurate solution with an approximation error of 1.51e-5 for the case with recursive utility and $\gamma = 10$. For the high persistence case with $\rho = 0.99$ a larger degree is required and the degree four polynomial is sufficient to compute a highly accurate solution. Overall we observe that the projection method provides highly accurate solutions for all specifications considered in this example. For a detailed comparison of the different methods see Chapter 4 of this thesis.

Table 1.5: Accuracy of the Projection Method

	Closed-Form	**Log-Lin**	**Projection**			**Discretization**		
			$n = 1$, $n_\sigma = 1$	$n = 4$, $n_\sigma = 4$	$n = 16$, $n_\sigma = 32$	$n_D = 5$	$n_D = 10$	$n_D = 50$
			$\boldsymbol{\psi = 1.5, \gamma = 1/\psi}$					
			$\rho = 0.95$					
$E(\frac{W}{C})$	1681.20	1681.16	1681.18	1681.20	1681.20	1669.99	1670.75	1671.00
Error	0	2.11e-5	1.19e-5	2.61e-8	2.60e-8	0.0067	0.0062	0.0060
			$\rho = 0.99$					
$E(\frac{W}{C})$	1868.36	1862.93	1865.54	1868.36	1868.36	3404.73	2121.64	1852.27
Error	0	0.0029	0.0015	1.21e-7	7.65e-11	0.8223	0.1356	0.0086
			$\boldsymbol{\psi = 1.5, \gamma = 10}$					
			$\rho = 0.95$					
$E(\frac{W}{C})$	-	1314.39	1314.59	1314.61	1314.61	1532.25	1514.25	1508.08
Error	-	1.66e-4	1.51e-5	4.37e-11	2.12e-12	0.1655	0.1518	0.1472
			$\rho = 0.99$					
$E(\frac{W}{C})$	-	517.13	518.97	529.39	529.39	869.23	653.99	570.65
Error	-	0.0231	0.0196	1.43e-9	4.91e-11	0.6419	0.2353	0.0779

The table shows the mean wealth-consumption ratio for the long-run risk model of Bansal and Yaron (2004) with constant volatility. Results are shown for the log-linearization, the projection as well as the discretization by Tauchen and Hussey (1991) with the extension of Floden (2007) that performs better for highly persistent processes. For the projection method solutions with three different degrees n where the approximation interval is set up n_σ standard deviations around the unconditional mean of the long-run risk process x_t are provided. For the discretization results are shown for three different numbers of approximation nodes n_D. The table also shows the relative error of the solutions, where in the case of $\gamma = 1/\psi$ the closed form solution is taken from de Groot (2015) and in the case of $\gamma \neq 1/\psi$ we compute the accurate solution by solving the model using the discretization method with a very large number of discretization nodes or equivalently the projection with a very large degree and state space. We use the same calibration as in Section 1.3.3 with $\delta = 0.9989$, $\mu_c = 0.0015$, $\bar{\sigma}^2 = 0.0078^2$ and $\phi_x = 0.044$.

1.C The Volatility of Volatility Factor

This section analyzes how log-linearization affects model outcomes when the model dynamics are described by a square-root process as for example in Bollerslev, Tauchen, and Zhou (2009), Tauchen (2011) or Bollerslev, Xu, and Zhou (2015). For this purpose we use the parsimonious model formulation as in Bollerslev, Tauchen, and Zhou (2009) who take the basic model setup (1.7) without the long-run risk factor $\phi_x = 0$ and add vol-of vol modeled by a square root process q_t:

$$\begin{aligned} \sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \sqrt{q_t}\eta_{\sigma,t+1} \\ q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\eta_{q,t+1} \\ \eta_{\sigma,t+1}, \eta_{q,t+1} &\sim i.i.d.\ N(0,1). \end{aligned} \tag{1.22}$$

As Tauchen (2011) notes, care is needed, as q_t can become negative in simulations if the volatility is too large compared to the mean of the process. The common approach in the literature is to assume a reflecting barrier at zero by replacing negative values with very small positive values to ensures positivity of the process (this approach has also been used for the stochastic volatility process in the original Bansal and Yaron (2004) study and many following papers). However, to compute model solutions, the assumption of a non-truncated distribution for the log-linearization is commonly used.

Take, for example, the calibration of Bollerslev, Tauchen, and Zhou (2009) given by $\delta = 0.997, \gamma = 10, \psi = 1.5, \mu_c = 0.0015, \nu = 0.978, \bar{\sigma}^2 = 0.0078^2$ and $\mu_q =$ 1e-6. Figure 1.11 shows model outcomes for CRRA preferences with $\psi = 1.5$ (Panel (a)) and the corresponding EZ case with $\gamma = 10$ (Panel (b)) for various persistence and volatility parameters of the vol-of-vol process ρ_q and ϕ_q. The black numbers show the true mean wealth-consumption ratio under the assumption of a reflecting boundary for q_t at zero. Blue values are the results from log-linearization under the assumption of a standard non-truncated normal distribution. Green circles denote convergence of both, the projection and the log-linearization approach. Red stars denote cases in which the log-linearization yields a complex solution, while the model solution using a truncated normal distribution is real. We find that, depending on the risk aversion, using the standard log-linearization technique can lead to complex solutions. This is for example the case for the calibration in Bollerslev, Tauchen, and Zhou (2009) with $\rho_q = 0.8$ and $\phi_q =$ 1e-3.[18]

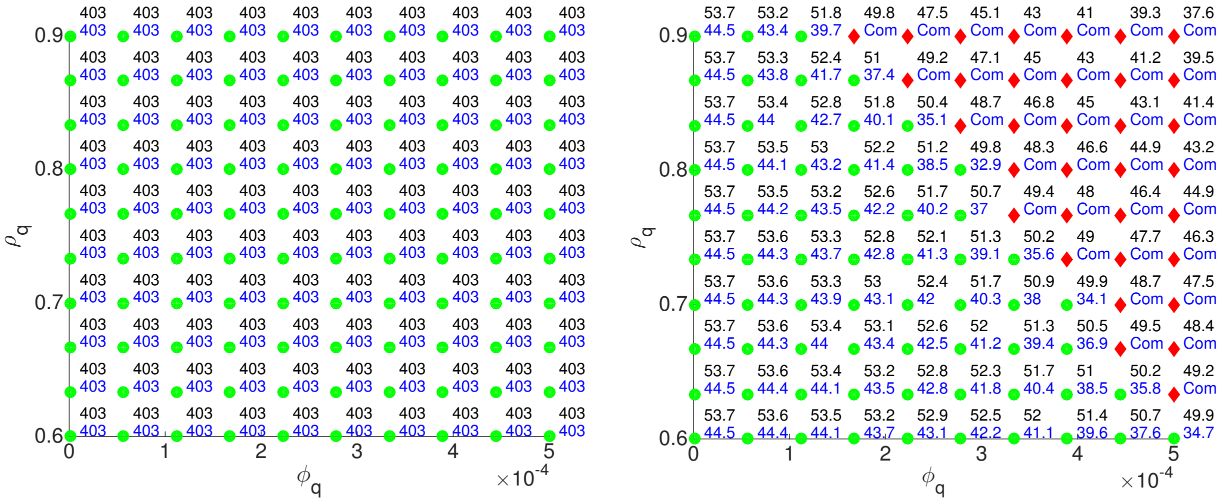
So what are the determinants of the complexity of the linearized solution? The square-root specification of q_t implies that the coefficient for q_t is determined by a quadratic equation and hence may have more than one solution. The log-linear approximation of the log wealth-consumption ratio $z_{w,t}$ has the following form

$$z_{w,t} = A_0 + A_\sigma \sigma_t^2 + A_q q_t \tag{1.23}$$

[18] Bollerslev, Tauchen, and Zhou (2009) provide a real solution by assuming a fixed value for the linearization constant $\kappa = 0.9$. However this approach doesn't give a solution to the model but ex ante fixes the mean value of the price dividend ratio and hence significantly biases the model outcome.

Figure 1.11: Sensitivity Analysis and Existence Results in the Vol-of-Vol Model

(a) $\boldsymbol{\psi = 1.5, \gamma = 2/3}$ (b) $\boldsymbol{\psi = 1.5, \gamma = 10}$



The graph shows the convergence properties as well as the mean wealth-consumption ratio for the vol-of-vol model of Bollerslev, Tauchen, and Zhou (2009). The results are reported for a range of persistence parameters ρ_q and volatility parameters ϕ_q. Panel (a) depicts the case of CRRA utility with $\psi = 1.5$, while panel (b) depicts the corresponding cases with recursive utility and $\gamma = 10$. Black numbers show the mean wealth-consumption ratio obtained by the projection approach using a reflecting barrier at zero and blue numbers show the values obtained by the standard log-linearization with normal shocks. Green circles denote convergence of both, the projection and the log-linearization approach. Red stars denote cases in which the log-linearization yields a complex solution, while the model solution using a truncated normal distribution is real. The model parameters are given by $\delta = 0.997, \mu_c = 0.0015, \nu = 0.978, \bar{\sigma}^2 = 0.0078^2$ and $\mu_q =$ 1e-6.

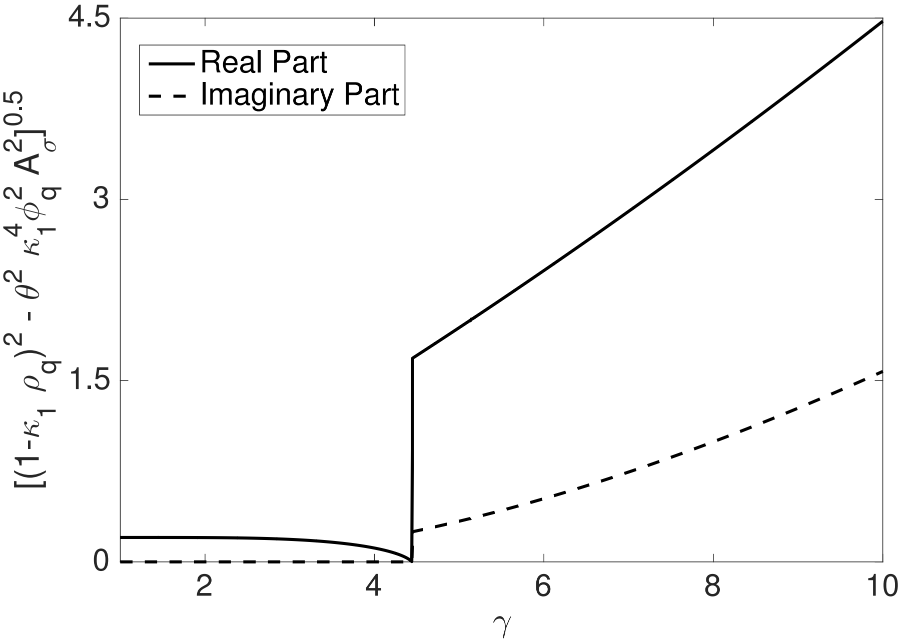
with the linearization coefficients given by[19]

$$\begin{aligned} A_\sigma &= \frac{(1-\gamma)^2}{2\theta(1-k_1\nu)} \\ A_0 &= \frac{\log\delta + (1-\frac{1}{\psi})\mu_c + k_0 + k_1\Big[A_\sigma\bar{\sigma}^2(1-\nu) + A_q\mu_q(1-\rho_q)\Big]}{(1-k_1)} \\ A_q &= \frac{1-k_1\rho_q \pm \sqrt{(1-k_1\rho_q)^2 - \theta^2 k_1^4\phi_q^2 A_\sigma^2}}{\theta k_1^2\phi_q^2}. \end{aligned} \tag{1.24}$$

We find that the coefficient for the vol-of-vol factor A_q has indeed two solutions. As Bollerslev, Tauchen, and Zhou (2009) show in their paper by the no arbitrage argument, the minus term is the economically meaningful root and the positive solution can be neglected. Complexity of the solution is determined by the term inside the square root in equation (1.24) given by $(1-k_1\rho_q)^2 - \theta^2 k_1^4\phi_q^2 A_\sigma^2$. So how does this term depend on the model parameters? Figure 1.12 shows the values of the square root term as a function of the risk aversion γ. In line with

[19] A brief description of how the linearization coefficients can be derived is shown in Appendix 4.A.2 of Chapter 4. A more general description is provided in Eraker (2008).

Figure 1.12: Analysis of Square-Root Term in the Vol-of-Vol Model



The graph shows the real and complex part of the square root term that determines A_q as a function of the risk aversion γ for the vol-of-vol model of Bollerslev, Tauchen, and Zhou (2009). The model parameters are given by $\delta = 0.997, \mu_c = 0.0015, \nu = 0.978, \bar{\sigma}^2 = 0.0078^2, \mu_q =$ 1e-6, $\rho_q = 0.8$ and $\phi_q = 1 \times 10^{-3}$.

the results above, we find that for small γ the solution is well behaved with only a real and no imaginary part. However if we increase γ, θ becomes significantly larger (it goes from -3 for $\gamma = 2$ to -27 for $\gamma = 10$) and hence the real part of the term decreases. For a certain threshold (about 4.4 in this example) the term hits zero and the solution thereafter consists of a significant imaginary part. Also Panel (b) in Figure 1.11 shows that the larger the persistence or the larger the volatility of the vol-of-vol process solutions become complex. Summarizing, using standard log-linearization with normal shocks to solve models with a large risk aversion and a persistent square-root process can yield complex solutions, even if real solutions under the assumption of a reflecting barrier exist. Hence when solving such models, either log-linearization with the assumption of a truncated normal distribution or more sophisticated methods like the projection approach described in this paper should be used.

Essay 2

Agent Heterogeneity and Asset Prices

Recursive Preferences, Agent Heterogeneity and Wealth Dynamics[1]

Karl Schmedders
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
karl.schmedders@business.uzh.ch

Walter Pohl
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
walter.pohl@business.uzh.ch

Ole Wilms
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
ole.wilms@business.uzh.ch

November 2015

Abstract

The paper shows how to solve asset pricing models with heterogeneous agents that have recursive utility. We derive first-order conditions to obtain optimal consumption shares and pricing functions. We present a methodology based on projection methods to solve for the equilibrium functions numerically and provide a simple example where we solve the long-run risk model of Bansal and Yaron (2004) with two agents that differ with regard to their risk aversion and intertemporal elasticity of substitution.

Keywords: Asset pricing; Long-run risk; Recursive preferences; Heterogeneous agents.

JEL Classification: G11, G12.

[1]We are indebted to Lars Hansen and Kenneth Judd for helpful discussions on the subject. We also thank Ferdinand Langnickel for editorial comments on the manuscript.

2.1 Introduction

In this paper we show how to solve asset pricing models in discrete time with heterogeneous agents that have recursive utility. We derive first-order conditions to obtain optimal consumption shares and pricing functions. We present a methodology based on projection methods to solve for the equilibrium functions numerically. Since our method is numerical, we are not constrained in the models we consider by analytical tractability. To demonstrate this, we provide a simple example where we solve the long-run risk model of Bansal and Yaron (2004) with two agents that differ with regard to their risk aversion and intertemporal elasticity of substitution. The long-run risk model, which is in many respects the leading consumption-based asset pricing model, has a complex exogenous process specification, and does not lend itself to non-numerical approaches.

Under the classical assumption of time-separable preferences, the influence of agent heterogeneity on market outcomes is well-understood. For example Sandroni (2000) or Blume and Easley (2006) have analyzed the influence of differences in beliefs on market selection. The market selection hypothesis, as first described by Alchian (1950) and Friedman (1953), states that agents with systematically wrong beliefs about the distribution of future quantities will lose wealth on average and will be eventually driven out of the market. So in the long-run only the agents with rational expectations will survive. Sandroni (2000) and Blume and Easley (2006) find strong support for this hypothesis under the assumption of time separable preferences. Judd, Kubler, and Schmedders (2003) show that trading volume is zero in a simple Lucas asset pricing model with heterogeneous agents and no growth. Agents trade once in the initial period and portfolio holdings are constant thereafter. The authors conclude that there must be other reasons than differences in agents with time separable preferences to explain the large trading volume observed in the data. Bhamra and Uppal (2014) show that in a simple endowment economy with growth and a common time preference rate, only the investor with the lowest degree of risk aversion survives in the long-run. So by market selection, even without differences in beliefs, the economy will eventually converge to a representative agent economy populated only by the least risk averse agent.

However, the past decades of research in asset pricing have shown that the assumption of time separable preferences leads to unrealistic degrees of risk aversion when it comes to explaining asset pricing data (e.g. Mehra and Prescott (1985) or Grossman and Shiller (1981)). Bansal and Yaron (2004) propose a solution combining recursive utility of Epstein-Zin type with highly persistent shocks to mean consumption growth. This approach has proven to be very successful in explaining long-standing asset pricing puzzles (e.g. Hansen, Heaton, and Li (2008), Bansal and Shaliastovich (2013), Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011) and Bansal, Kiku, and Yaron (2012a)). A key step for the success of the model is a preference structure that allows for the separation of risk aversion and the elasticity of intertemporal substitution (EIS) to model the preference for the early resolution of risk. This can not be achieved in a standard time separable utility model where the EIS equals the inverse of the risk aversion and hence there is neutrality about the timing of the resolution of risk.

A natural question that arises is whether the conclusions on agent heterogeneity drawn under the assumption of time separable preferences carry over to the generalized class of recursive utility. Recent research suggests that they do not. For example in contrast to the results in Bhamra and Uppal (2014), under general recursive preferences, multiple agents with differences in the utility parameters can survive even if there is growth (Branger, Dumitrescu, Ivanova, and Schlag (2011)). A difference in the degree of risk aversion can be offset by a difference in the intertemporal elasticity of substitution resulting in a range of utility parameter combinations where multiple agents survive. Also Borovička (2015) shows that agents with fundamentally wrong beliefs can survive or even dominate in an economy with recursive utility. So the inferences about market selection and equilibrium outcomes fundamentally differ under the assumption of general recursive utility compared to the special case of standard time separable preferences.

This paper presents a general framework to analyze the influence of agent heterogeneity in asset pricing models with recursive preferences. Opposed to the work of Branger, Dumitrescu, Ivanova, and Schlag (2011), Bhamra and Uppal (2014) and Borovička (2015), who consider models in continuous time, the methodology in this paper focuses on discrete time economies. The methodology builds on the work of Judd, Kubler, and Schmedders (2003) who show how to compute equilibria in a pure infinite-horizon exchange economy with time separable preferences. We extend their approach to feature general recursive utility, continuous state variables and growth. The generality of our approach makes it applicable to a large class of asset pricing models. In this paper we focus on heterogeneity with regard to the utility specification. However, it is straight-forward to generalize the approach to heterogeneity in beliefs.

The main contribution of the paper is to derive first-order conditions that describe the equilibrium. We present a computational approach, based on projection methods, to solve for the equilibrium quantities numerically and apply it to solve the long-run risk model of Bansal and Yaron (2004) with two agents that differ with regard to their preference parameters. While an extensive qualitative and quantitative analysis of the model is beyond the scope of this paper, it serves as an example to demonstrate the solution approach. The results give the reader insights about the influence of agent heterogeneity on equilibrium outcomes and motivate future research in the area. Interesting research questions that can be analyzed using the methodology from this paper are for example: How is the wealth distribution influenced by differences in the utility functions of the agents? How do different risks influence the wealth distribution? Can multiple agents with different utility specifications survive in the model and if yes, what are the mechanisms of survival? Or what are the asset pricing implications of agent heterogeneity?

The paper is organized as follows. Section 2.2 describes the methodology to solve asset pricing models with heterogeneous agents and recursive preferences. For this we first derive the first-order conditions describing the equilibrium for the general form of recursive preferences. Afterwards, in Section 2.2.1, we provide the specific expressions for the special case of Epstein-Zin preferences. As in general there are no closed-form solutions for the model, we present a computational procedure to solve for the equilibrium functions in Section 2.2.2. Section 2.2.3 shows how to exploit properties of the value function to obtain more accurate and robust approximations. In Section 2.3 we apply the method to solve the long-run risk model of Bansal and Yaron (2004) with two investors and Section 2.4 concludes.

2.2 Solution Method

We consider a standard Lucas asset pricing model with heterogeneous agents and complete markets. Time is discrete and indexed by $t \in \mathbb{N}_0 \equiv \{0, 1, \ldots\}$. Let $\boldsymbol{s_t} \in \mathbb{R}^l$, $l \geq 1$ denote the state of the economy which is populated by a finite number of infinitely-lived agents of types $\mathbb{H} = \{1, 2, \ldots H\}$. We also define $\mathbb{H}^- = \{2, 3, \ldots H\}$. Aggregate consumption $C(\boldsymbol{s_t})$ is exogenous and only depends on the state of the economy. Let $c^h(\boldsymbol{s_t})$ denote individual consumption in period t and $\{c^h\}_t = \{c^h(\boldsymbol{s_t}), c^h(\boldsymbol{s_{t+1}}), \ldots\}$ denotes the consumption stream of agent $h = 1, \ldots H$. Market clearing requires that

$$\sum_{h=1}^{H} c^h(\boldsymbol{s_t}) = C(\boldsymbol{s_t}). \tag{2.1}$$

Each agent has recursive utility $U^h(\{c^h\}_t)$ specified by an aggregator $F^h(c, x)$ and a certainty-equivalence function $G(x)$:

$$U^h(\{c^h\}_t) = F^h\left(c^h(\boldsymbol{s_t}), R_t^h\left[U^h(\{c^h\}_{t+1})\right]\right) \tag{2.2}$$

where

$$R_t^h(x_{t+1}) = G_h^{-1}(E_t[G_h(x_{t+1})]). \tag{2.3}$$

To solve the model we write it as a social planner's problem. The social planner maximizes a weighted sum of the individual agent's utilities in $t = 0$:

$$SP(\{\boldsymbol{c}\}_0, \boldsymbol{\lambda}) = \sum_{h=1}^{H} \lambda^h U^h\left(\{c^h\}_0\right) \tag{2.4}$$

where each agent has a weight $\boldsymbol{\lambda} = \{\lambda^1, \ldots, \lambda^H\}$, known as the Negishi weight $\{\boldsymbol{c}\}_0 = \{\{c^1\}_0, \ldots, \{c^H\}_0\}$. The optimal decision of the social planner in the initial period takes into account all future consumption streams of the individual agents and the optimal decisions must satisfy the market clearing condition (2.1). For the ease of notation we abbreviate the state dependence in the following, so we use c_t^h for $c^h(\boldsymbol{s_t})$ and $U^h_{\{t\}}$ for $U^h\left(\{c^h\}_t\right)$.

To derive the first-order conditions, we borrow a technique from the calculus of variations. For any function f_t we can vary the consumption of two agents by

$$\begin{aligned} c_t^h &\rightarrow c_t^h + \epsilon f_t \\ c_t^l &\rightarrow c_t^l - \epsilon f_t. \end{aligned} \tag{2.5}$$

It is sufficient to consider the variation with agent $l = 1$. Since we have an optimal allocation it must be true that

$$\left.\frac{\mathrm{d}SP(\{\boldsymbol{c}\}_0, \boldsymbol{\lambda})}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = 0. \tag{2.6}$$

This gives us

$$\lambda^h \hat{U}^h_{0,t} = \lambda^1 \hat{U}^1_{0,t}, \qquad h \in \mathbb{H}^-, \tag{2.7}$$

where $\hat{U}^h_{t,t+k}$ is defined as

$$\hat{U}^h_{t,t+k} = \left.\frac{\mathrm{d}U^h(c_t^h, \ldots, c_{t+k}^h + \epsilon f_{t+k}, \ldots)}{\mathrm{d}\epsilon}\right|_{\epsilon=0}. \tag{2.8}$$

We normalize λ^h by $\underline{\lambda}^h \equiv \frac{\lambda^h}{\lambda^1 + \lambda^h} \in (0,1)$ and obtain[2]

$$\underline{\lambda}^h \hat{U}^h_{0,t} = (1 - \underline{\lambda}^h) \hat{U}^1_{0,t} \qquad h \in \mathbb{H}^-. \tag{2.9}$$

$\hat{U}^h_{t,t+k}$ satisfies a recursive equation with the initial condition

$$\hat{U}^h_{t,t} = \left.\frac{\mathrm{d}U^h(c_t^h + \epsilon f_t, \ldots)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = F_1^h\left(c_t^h, R_t[U^h_{\{t+1\}}]\right) \cdot f_t \tag{2.10}$$

where $F_k^h\left(c_t^h, R_t^h[U^h_{\{t+1\}}]\right)$ denotes the derivative of $F^h\left(c_t^h, R_t^h[U^h_{\{t+1\}}]\right)$ with respect to its kth argument. The recursive step is given by

[2] Another approach that yields equivalent results is to assume that $\lambda^1 = 1$. This implies that $\underline{\lambda}^h \in \mathbb{R}^+$. However, for computational reasons it is more appealing to have $\underline{\lambda}^h \in (0,1)$.

$$
\begin{aligned}
\hat{U}^h_{t,t+k} &= \left.\frac{\mathrm{d}F^h\left(c^h_t, R^h_t\left[U^h(c^h_{t+1},\ldots c^h_{t+k}+\epsilon f_{t+k},\ldots)\right]\right)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\left.\frac{\mathrm{d}R^h_t\left[U^h(\cdot)\right]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{\mathrm{d}G^{-1}_h\left(E_tG_h\left[U^h(\cdot)\right]\right)}{\mathrm{d}E_tG_h[U^h(\cdot)]}\cdot\left.\frac{\mathrm{d}E_tG_h[U^h(\cdot)]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{1}{G'_h(G^{-1}_h(E_tG_h[U^h_{\{t+1\}}]))}\cdot E_t\left(G'_h(U^h_{\{t+1\}})\cdot\hat{U}^h_{t+1,t+k}\right) \\
&= F^h_2\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{E_t\left(G'_h(U^h_{\{t+1\}})\cdot\hat{U}^h_{t+1,t+k}\right)}{G'_h(R^h_t[U^h_{\{t+1\}}])}
\end{aligned}
\tag{2.11}
$$

where we use $\frac{\partial G^{-1}(x)}{\partial x}=\frac{1}{G'(G^{-1}(x))}$ and abbreviate $U^h(c^h_{t+1},\ldots c^h_{t+k}+\epsilon f_{t+k},\ldots)$ by $U^h(\cdot)$. We can recast this recursion into a useful form. Therefore we define a second recursion $U^h_{t,t+k}$ by

$$U^h_{t,t}=F^h_1\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right) \tag{2.12}$$

and

$$U^h_{t,t+k}=\Pi^h_{t+1}\cdot U^h_{t+1,t+k} \tag{2.13}$$

where

$$\Pi^h_{t+1}=F^h_2\left(c^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{G'_h(U^h_{\{t+1\}})}{G'_h(R^h_t[U^h_{\{t+1\}}])}. \tag{2.14}$$

A simple induction shows that

$$\hat{U}^h_{t,t+k}=E_t(U^h_{t,t+k}f_t). \tag{2.15}$$

Plugging (2.15) into the optimality condition (2.9) we get

$$E_0((\underline{\lambda}^hU^h_{0,t}-(1-\underline{\lambda}^h)U^1_{0,t})f_t)=0, \qquad h\in\mathbb{H}^-. \tag{2.16}$$

The expression above has to hold for any function f_t. Hence, by the fundamental lemma of the calculus of variations (Gelfand and Fomin (1963)), it must be true that

$$\underline{\lambda}^hU^h_{0,t}=(1-\underline{\lambda}^h)U^1_{0,t}, \qquad h\in\mathbb{H}^-. \tag{2.17}$$

We can then split expression (2.17) into two parts. First define $\underline{\lambda}_0^h \equiv \underline{\lambda}^h$ to obtain

$$\frac{\underline{\lambda}_0^h}{1-\underline{\lambda}_0^h} = \frac{U_{0,t}^1}{U_{0,t}^h} = \frac{\Pi_0^1}{\Pi_0^h}\frac{U_{1,t}^1}{U_{1,t}^h} = \frac{\Pi_0^1}{\Pi_0^h}\frac{\underline{\lambda}_1^h}{1-\underline{\lambda}_1^h}, \qquad h \in \mathbb{H}^-,$$

where $\underline{\lambda}_1^h$ denotes the Negishi weight of the social planner's optimum in $t = 1$. Generalizing this equation for any period t, we obtain the following dynamics for the optimal weight $\underline{\lambda}_{t+1}^h$

$$\frac{\underline{\lambda}_t^h}{1-\underline{\lambda}_t^h} = \frac{\Pi_{t+1}^1}{\Pi_{t+1}^h}\frac{\underline{\lambda}_{t+1}^h}{1-\underline{\lambda}_{t+1}^h}, \qquad h \in \mathbb{H}^-$$

or equivalently

$$\underline{\lambda}_{t+1}^h = \frac{\Pi_{t+1}^h \underline{\lambda}_t^h}{(1-\underline{\lambda}_t^h)\Pi_{t+1}^1 + \underline{\lambda}_t^h \Pi_{t+1}^h}, \qquad h \in \mathbb{H}^-. \tag{2.18}$$

The second expression is obtained by inserting the initial condition (2.12) into (2.17) for $t = 0$ and generalizing it for any social planner's optimum at time t:

$$\underline{\lambda}_t^h F_1^h\left(c_t^h, R_t^h[U_{\{t+1\}}^h]\right) = (1-\underline{\lambda}_t^h) F_1^1\left(c_t^1, R_t^1[U_{\{t+1\}}^1]\right), \qquad h \in \mathbb{H}^-. \tag{2.19}$$

Equation (2.19) states the optimality conditions for the individual consumption choices at any time t. Note that for time separable utility, $F_1^h\left(c_t^h, R_t^h[U_{\{t+1\}}^h]\right)$ is simply marginal utility of agent h at time t, and so we obtain the same optimality condition as for example in Judd, Kubler, and Schmedders (2003) (compare equation (7) on page 2209).[3] In this special case the Negishi weights can be pinned down in the initial period and thereafter remain constant. For general recursive preferences this is not true. The optimal weights vary over time following the law of motion described by equation (2.18).

We can use the two equations (2.18) and (2.19) together with the market clearing condition (2.1) to compute the social planner's optimum. We therefore define $\boldsymbol{\underline{\lambda}_t^-} = \{\underline{\lambda}_t^2, \underline{\lambda}_t^3, \ldots, \underline{\lambda}_t^H\}$ and let V^h denote the value function of agent $h \in \mathbb{H}$. We are looking for model solutions of the form $V^h(\boldsymbol{\underline{\lambda}_t^-}, \boldsymbol{s_t})$. So additional to the exogenous states $\boldsymbol{s_t}$, the model solution depends on the time varying Negishi weights $\boldsymbol{\underline{\lambda}_t^-} \in (0,1)^{H-1}$. An optimal allocation is then characterized by the following four equations:

[3] Judd, Kubler, and Schmedders (2003) set $\lambda^1 = 1$ instead of normalizing the weight to the interval $(0, 1)$.

- the market clearing condition (2.1)

$$\sum_{h=1}^{H} c^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}) = C(\boldsymbol{s_t}) \tag{2.20}$$

- the optimality conditions (2.19) for the individual consumption decisions

$$\underline{\lambda}_t^h F_1^h \left(c^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}), R_t^h[V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}})]\right) = (1-\underline{\lambda}_t^h) F_1^1 \left(c^1(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}), R_t^1[V^1(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}})]\right) \tag{2.21}$$

for $h \in \mathbb{H}^-$

- the value functions (2.2) of the individual agents

$$V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}) = F^h\left(c^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}), R_t^h[V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}})]\right), \qquad h \in \mathbb{H} \tag{2.22}$$

- the equations (2.18) for the dynamics of $\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-$

$$\underline{\lambda}_{t+1}^h = \frac{\Pi_{t+1}^h \underline{\lambda}_t^h}{(1-\underline{\lambda}_t^h)\Pi_{t+1}^1 + \underline{\lambda}_t^h \Pi_{t+1}^h}, \qquad h \in \mathbb{H}^-, \tag{2.23}$$

where

$$\Pi_{t+1}^h = F_2^h\left(c^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t}), R_t^h[V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}})]\right) \cdot \frac{G_h'(V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}}))}{G_h'(R_t^h[V^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t+1}}^-, \boldsymbol{s_{t+1}})])} \tag{2.24}$$

This concludes the general description of the equilibrium obtained from the social planer's optimization problem.

2.2.1 The Case of Epstein-Zin Preferences

In this section we provide the specific expressions for V^h, F_1^h, F_2^h and Π^h when the heterogeneous investors have recursive preferences as in Epstein and Zin (1989) and Weil (1989). The value function for Epstein-Zin (EZ) preferences is usually expressed as[4]

$$V_t^{h,EZ} = \left[(1-\delta^h)(c_t^h)^{1-\frac{1}{\psi^h}} + \delta^h \left[E_t\left((V_{t+1}^{h,EZ})^{1-\gamma^h}\right)\right]^{\frac{1-\frac{1}{\psi^h}}{1-\gamma^h}}\right]^{\frac{1}{1-\frac{1}{\psi^h}}}.$$

[4]For the ease of notation we again abbreviate the dependence on the exogenous state $\boldsymbol{s_t}$ and the endogenous state $\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-$. Hence we write $V_t^{h,EZ}$ for $V^{h,EZ}(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t})$ or c_t^h for $c^h(\underline{\boldsymbol{\lambda}}_{\boldsymbol{t}}^-, \boldsymbol{s_t})$.

For the purpose of this paper it is more useful to use $V_t^h \equiv \frac{(V_t^{h,EZ})^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}}$ that allows for easier derivatives as shown in the following lines. We then have

$$V_t^h = (1-\delta^h)\frac{(c_t^h)^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}} + \delta^h R_t^h\left[V_{t+1}^h\right] \tag{2.25}$$

where

$$\begin{aligned} R_t^h\left[V_{t+1}^h\right] &= G_h^{-1}\left(E_t\left[G_h(V_{t+1}^h)\right]\right) \\ G_h(V_{t+1}^h) &= \left(\left(1-\frac{1}{\psi^h}\right)V_{t+1}^h\right)^{\frac{1-\gamma^h}{1-\frac{1}{\psi^h}}}. \end{aligned}$$

Using this notation the derivatives of $F^h\left(c_t^h, R_t^h[U_{t+1}^h]\right) = V_t^h$ with respect to its first and second argument are then given by

$$F_{1,t}^h = (1-\delta^h)(c_t^h)^{-\frac{1}{\psi^h}} \tag{2.26}$$

and

$$F_{2,t}^h = \delta^h. \tag{2.27}$$

Also we have

$$G_h'(V_{t+1}^h) = (1-\gamma^h)\left(\left(1-\frac{1}{\psi^h}\right)V_{t+1}^h\right)^{\frac{\frac{1}{\psi^h}-\gamma^h}{1-\frac{1}{\psi^h}}},$$

and for Π_{t+1}^h we obtain (see equation (2.24))

$$\Pi_{t+1}^h = \delta^h\left(\frac{V_{t+1}^h}{R_t^h\left[V_{t+1}^h\right]}\right)^{\frac{\frac{1}{\psi^h}-\gamma^h}{1-\frac{1}{\psi^h}}}. \tag{2.28}$$

Plugging the specific expressions for EZ preferences into the general first-order conditions (2.20)-(2.24) we obtain the following system that describes the equilibrium:

The market clearing condition:

$$\sum_{h=1}^{H} c_t^h = C(\boldsymbol{s_t}). \tag{MC}$$

The optimality condition for the individual consumption decisions:

$$\underline{\lambda}_t^h (1-\delta^h)(c_t^h)^{-\frac{1}{\psi^h}} = (1-\underline{\lambda}_t^h)(1-\delta^1)(c_t^1)^{-\frac{1}{\psi^1}}, \quad h \in \mathbb{H}^-. \tag{CD}$$

The value functions of the individual agents:

$$V_t^h = (1-\delta^h)\frac{(c_t^h)^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}} + \delta^h R_t^h \left[V_{t+1}^h\right], \qquad h \in \mathbb{H}. \tag{VF}$$

The equation for the dynamics of $\boldsymbol{\underline{\lambda}_t^-}$:

$$\begin{aligned} \underline{\lambda}_{t+1}^h &= \frac{\Pi_{t+1}^h \underline{\lambda}_t^h}{(1-\underline{\lambda}_t^h)\Pi_{t+1}^1 + \underline{\lambda}_t^h \Pi_{t+1}^h} \\ \Pi_{t+1}^h &= \delta^h \left(\frac{V_{t+1}^h}{R_t^h\left[V_{t+1}^h\right]}\right)^{\frac{\frac{1}{\psi^h}-\gamma^h}{1-\frac{1}{\psi^h}}}, \qquad h \in \mathbb{H}^-. \end{aligned} \tag{Dλ}$$

Note that equation (CD) and hence the individual consumption decisions c_t^h only depend on time t information and there is no intertemporal dependence. This allows us to first solve for c_t^h given the current state of the economy and in a second step solve for the dynamics of the Negishi weights. Hence, we can separate solving the optimality conditions (MC)-(Dλ) into two steps in order to reduce the computational complexity. In the following section we describe this approach in detail.

2.2.2 Computational Procedure - A Two Step Approach

For the ease of notation the following procedures are described for $H = 2$ agents and a single state variable $s_t \in \mathbb{R}^1$. However, the approach can analogously be extended to the general case of $H > 2$ agents and multiple states. We solve the social planner's problem using a collocation projection.[5] For this we perform the usual transformation from an equilibrium described by

[5] We provide a detailed description of projection methods and how they can be applied to solve asset pricing models with recursive preferences in Chapter 4 of this thesis. In this paper we describe how the method can be used to solve the heterogeneous agent economy and we assume that the general projection approach from Chapter 4 is known to the reader.

the infinite sequences (with a time index t) to the equilibrium being described by functions of some state variable(s) x on a state space X. We denote the current exogenous state of the economy by s and the subsequent state in the next period by s' with the state space $S \in \mathbb{R}^1$. $\underline{\lambda}_2$ denotes the current endogenous state of the Negishi weight and $\underline{\lambda}_2'$ denotes the corresponding state in the subsequent period with $\underline{\Lambda}_2 \in (0,1)$.

We approximate the value functions of the two agents $V^h(\underline{\lambda}_2, s), h = \{1,2\}$ by a set of Chebychev basis functions $\{\Upsilon_i\}_{i\in\{0,1,\ldots,n\}}$:

$$\hat{V}^h(\underline{\lambda}_2, s; \boldsymbol{\alpha^h}) = \sum_{i=0}^{n}\sum_{j=0}^{m} \alpha_{i,j}^h \Upsilon_i(\underline{\lambda}_2)\Upsilon_j(s), \qquad h = \{1,2\} \tag{2.29}$$

where $\boldsymbol{\alpha}^h$ are $n \times m$ matrices of unknown coefficients. For the collocation projection we have to choose a set of collocation nodes $\{\underline{\lambda}_{2_k}\}_{k=0}^{n}$ and $\{s_l\}_{l=0}^{m}$ at which we evaluate $\hat{V}^h(\underline{\lambda}_2, s; \boldsymbol{\alpha^h})$. In the following we show how to first solve for the individual consumption levels at the collocation nodes $c_{k,l}^h = c^h(\underline{\lambda}_{2_k}, s_l)$ that are then used to solve for the value functions V^h and the dynamics of the endogenous state $\underline{\lambda}_2$.

Step 1: *Computing Optimal Consumption Allocations*

Equation (CD) has to hold at each collocation node $\{\underline{\lambda}_{2_k}, s_l\}_{k=0,l=0}^{n,m}$:

$$\underline{\lambda}_{2_k}(1-\delta^2)\left(c_{k,l}^2\right)^{-\frac{1}{\psi^2}} = (1-\underline{\lambda}_{2_k})(1-\delta^1)\left(c_{k,l}^1\right)^{-\frac{1}{\psi^1}}.$$

Together with the market clearing condition (MC) we get

$$\underline{\lambda}_{2_k}(1-\delta^2)\left(c_{k,l}^2\right)^{-\frac{1}{\psi^2}} = (1-\underline{\lambda}_{2_k})(1-\delta^1)\left(C(s_l)-c_{k,l}^2\right)^{-\frac{1}{\psi^1}}. \tag{2.30}$$

So for each node $\{\underline{\lambda}_{2_k}, s_l\}_{k=0,l=0}^{n,m}$ the optimal consumption choice $c_{k,l}^2$ can be computed by solving equation (2.30) and $c_{k,l}^1$ is obtained by the market clearing condition (MC).[6]

[6]Note that in the case of H agents we have to solve a system of $H-1$ equations that pin down the $H-1$ individual consumption choices $c^h \in \mathbb{H}^-$.

Step 2: *Solving for the Value Function and the Dynamics of the Negishi Weights*

Solving for the value function is not straight-forward as it depends on the dynamics of the endogenous state $\underline{\lambda}_2$ that are unknown and follow equation (Dλ). We compute the expectation over the exogenous state by a Gauss-Quadrature with Q quadrature nodes. This implies that the values for s' at which we evaluate V^h are given by the quadrature rule. We denote the corresponding quadrature nodes by $\{s'_{l,g}\}_{l=0,g=1}^{m,Q}$ and the weights by $\{\omega_g\}_{g=1}^{Q}$.[7] We can then solve equation (Dλ) for a given pair of collocation nodes $\{\underline{\lambda}_{2_k}, s_l\}_{k=0,l=0}^{n,m}$ and the corresponding quadrature nodes $\{s'_{l,g}\}_{l=0,g=1}^{m,G}$ to compute a vector $\vec{\underline{\lambda}}'_2$ of size $(n+1)\times(m+1)\times G$ that consists of the corresponding values $\underline{\lambda}'_{2_{k,l,g}}$ for each node. For each $\underline{\lambda}'_{2_{k,l,g}}$ equation (Dλ) then reads

$$\begin{aligned} \underline{\lambda}'_{2_{k,l,g}} &= \frac{\underline{\lambda}_{2_k}\Pi^2}{(1-\underline{\lambda}_{2_k})\Pi^1 + \underline{\lambda}_{2_k}\Pi^2} \\ \Pi^h &= \delta^h \left(\frac{V^h(\underline{\lambda}'_{2_{k,l,g}}, s'_{l,g})}{R^h\left[V^h(\underline{\lambda}'_2, s')|\underline{\lambda}_{2_k}, s_l\right]} \right)^{\frac{\frac{1}{\psi^h}-\gamma^h}{1-\frac{1}{\psi^h}}} \end{aligned} \tag{2.31}$$

where

$$R^h\left[V^h(\underline{\lambda}'_2, s')|\underline{\lambda}_{2_k}, s_l\right] = G_h^{-1}\left(E\left[G_h(V^h(\underline{\lambda}'_2, s'))|\underline{\lambda}_{2_k}, s_l\right]\right).$$

Note that $\underline{\lambda}'_{2_{k,l,g}}$ still depends on the full distribution of $\underline{\lambda}'_2$ through the expectation operator. By applying the Gauss-Quadrature to compute the expectation we get

$$E\left[G_h(V^h(\underline{\lambda}'_2, s'))|\underline{\lambda}_{2_k}, s_l\right] \approx \sum_{g=1}^{G} G_h(V^h(\underline{\lambda}'_{2_{k,l,g}}, s^Q_{l,g}))\cdot\omega_g.$$

So by computing the expectation with the quadrature rule, we do not need the full distribution of $\underline{\lambda}'_2$ but only have to evaluate V^h at the values $\underline{\lambda}'_{2_{k,l,g}}$ that can be obtained by solving (2.31) for each pair of collocation nodes $\{\underline{\lambda}_{2_k}, s_l\}_{k=0,l=0}^{n,m}$ and the corresponding quadrature nodes $\{s'_{l,g}\}_{l=0,g=1}^{m,G}$. So at the end we have a square system of equations with $(n+1)\times(m+1)\times G$ unknowns $\underline{\lambda}'_{2_{k,l,g}}$ and as many equations (2.31) for each $\{k,l,g\}$.

The value function is in general not known so we have to compute it simultaneously when solving for $\underline{\lambda}'_{2_{k,l,g}}$. Plugging the approximation (2.29) into the value function (VF) yields

$$\hat{V}^h(\underline{\lambda}_{2_k}, s_l; \boldsymbol{\alpha^h}) = (1-\delta^h)\frac{\left(c^h_{k,l}\right)^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}} + \delta^h R^h\left[\hat{V}^h(\underline{\lambda}'_2, s'; \boldsymbol{\alpha^h})|\underline{\lambda}_{2_k}, s_l\right]. \tag{2.32}$$

[7] Note that the quadrature nodes $\{\{s'_{l,g}\}_{g=0}^{G}\}_{l=0}^{m}$ depend on the state today $\{s_l\}_{l=0}^{m}$.

The collocation projection conditions require that the equation has to hold at each collocation node $\{\underline{\lambda}_{2_k}, s_l\}_{k=0,l=0}^{n,m}$. So we obtain a system of equations with $(n+1)\times(m+1)\times 2$ unknown coefficients $\boldsymbol{\alpha^h}, h \in \{1,2\}$ and as many equations (2.32) for each collocation node that we solve simultaneously with the system for $\underline{\lambda}'_{2_{k,l,g}}$ described above.

2.2.3 Properties of the Value Function

In the case of heterogeneous agents the approximation of the value function is a delicate computational task as an agent can die out over time. Marginal utility of the agent at this limiting case is infinity which makes it difficult to obtain accurate approximations for the value function close to the singularity. To obtain information about the properties of the singularity, we formally derive the limiting behavior of the value function for the special case of an economy with no uncertainty. We then include this information in the value function approximation for the stochastic economy. This leads to significant improvements in the accuracy of the approximation of the value function as we demonstrate thereafter.

From equation (CD) we know that

$$c^2(\underline{\lambda}_2, s) = \left(\frac{1-\delta^1}{1-\delta^2}\right)^{\psi^2} (\underline{\lambda}_2)^{\psi^2}(1-\underline{\lambda}_2)^{-\psi^2}(c^1(\underline{\lambda}_2, s))^{\frac{\psi^2}{\psi^1}}. \tag{2.33}$$

We are interested in the properties of $c^2(\underline{\lambda}_2, s)$ for $\underline{\lambda}_2$ close to 0. For $\underline{\lambda}_2 \approx 0$ agent 1 obtains all consumption so $c^1(\underline{\lambda}_2, s) \approx C(s)$ and the Negishi weight of the first agent becomes 1. Therefore we obtain

$$c^2(\underline{\lambda}_2, s) \approx \left(\frac{1-\delta^1}{1-\delta^2}\right)^{\psi^2} (\underline{\lambda}_2)^{\psi^2} C(s)^{\frac{\psi^2}{\psi^1}} \tag{2.34}$$

for $\underline{\lambda}_2$ close to 0. The value function (2.25) for the deterministic economy at the steady state $s = s', \underline{\lambda}_2 = \underline{\lambda}'_2$ is given by

$$V^h(\underline{\lambda}_2, s) = \frac{(c^h(\underline{\lambda}_2, s))^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}}. \tag{2.35}$$

Inserting the behavior of $c^2(\underline{\lambda}_2, s)$ for $\underline{\lambda}_2$ close to 0, we obtain[8]

$$V^2(\underline{\lambda}_2, s) \approx \frac{\left(\frac{1-\delta^1}{1-\delta^2}\right)^{\psi^2-1} (\underline{\lambda}_2)^{\psi^2-1} C(s)^{\frac{\psi^2-1}{\psi^1}}}{1-\frac{1}{\psi^2}} \equiv \Upsilon^0(\underline{\lambda}_2, s). \tag{2.36}$$

[8] For the first agent we obtain a similar expression for $\underline{\lambda}_2$ close to 1 given by $V^1(\underline{\lambda}_2, s) \approx \frac{\left(\frac{1-\delta^2}{1-\delta^1}\right)^{\psi^1-1}(1-\underline{\lambda}_2)^{\psi^1-1}C(s)^{\frac{\psi^1-1}{\psi^2}}}{1-\frac{1}{\psi^1}}$.

We denote by $\Upsilon^0(\underline{\lambda}_2, s)$ the zero basis functions which we add to the value function approximation (2.29) to obtain accurate approximations close to the singularity:

$$\hat{V^2}(\underline{\lambda}_2, s; \boldsymbol{\alpha^2}) = \sum_{i=0}^{n}\sum_{j=0}^{m} \alpha^2_{i,j}\Upsilon_i(\underline{\lambda}_2)\Upsilon_j(s) + \Upsilon^0(\underline{\lambda}_2, s). \tag{2.37}$$

In the following we provide an example how including the zero basis function can improve the polynomial approximation of the value function.

Consider a deterministic economy with two agents that differ with respect to their EIS. Aggregate consumption is given by $C(s) = 1, \forall s$ and assume for the moment that $\underline{\lambda}_2$ is constant over time. The value function of agent h for a given $\underline{\lambda}_2$ is then given by

$$V^h(\underline{\lambda}_2) = \frac{\left(c^h(\underline{\lambda}_2)\right)^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}} \tag{2.38}$$

where $c^2(\underline{\lambda}_2)$ is obtained by solving

$$\underline{\lambda}_2\left(c^2(\underline{\lambda}_2)\right)^{-\frac{1}{\psi^2}} = (1-\underline{\lambda}_2)\left(1-c^2(\underline{\lambda}_2)\right)^{-\frac{1}{\psi^1}} \tag{2.39}$$

and $c^1(\underline{\lambda}_2) = 1 - c^2(\underline{\lambda}_2)$. In Figure 2.1 we compare the accuracy of approximating the value function with and without the zero basis functions for $\psi^1 = 1.5$ and $\psi^1 = 1.2$.

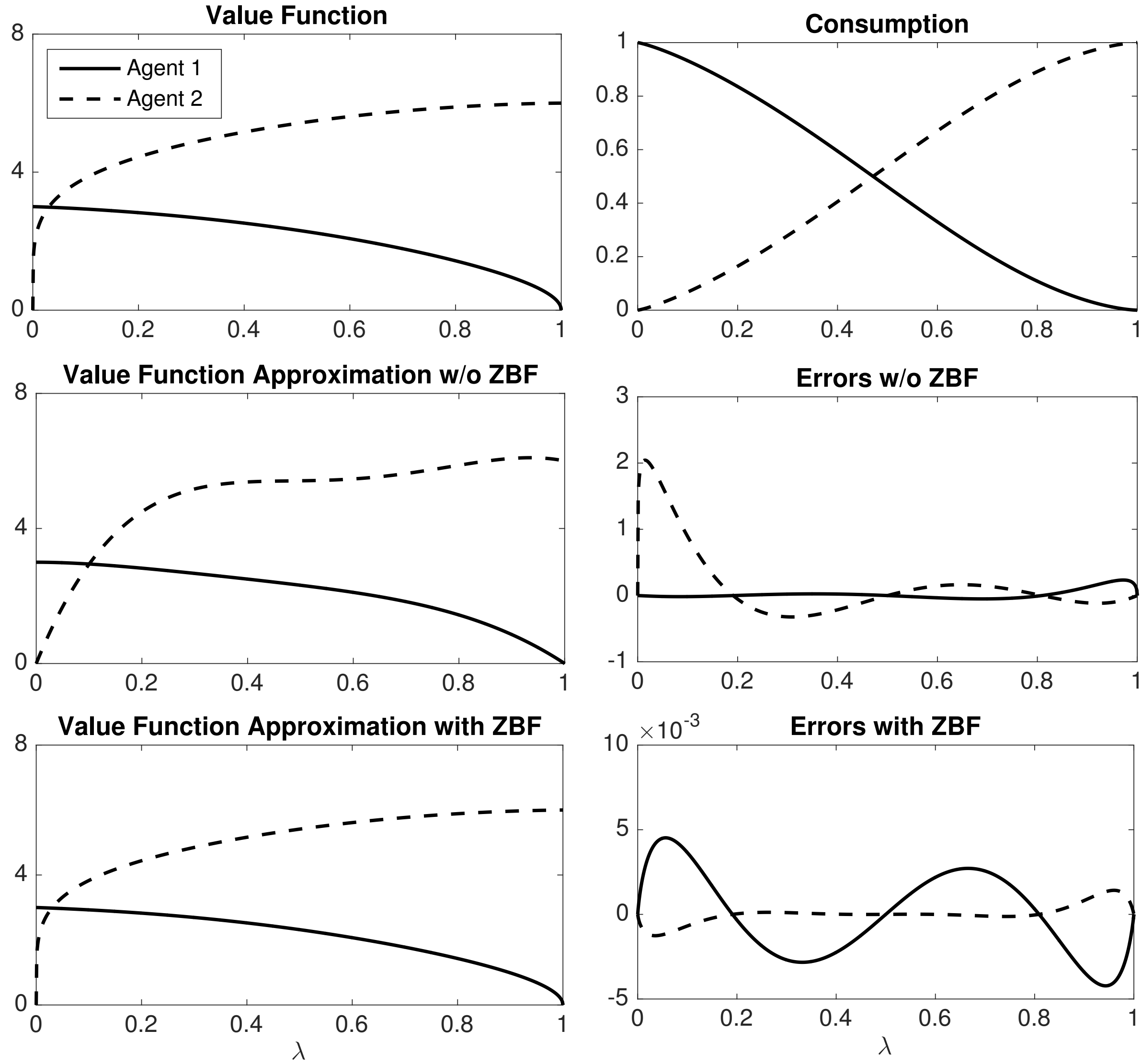
The first row shows closed-form solution of the value function (2.38) as well as the consumption choices of the two agents obtained from (2.39) for $\underline{\lambda}_2 \in (0, 1)$. We find that the value function increases strongly close to the singularities of the two agents that might introduce difficulties when polynomial approximations are used. In the second row we show the value function approximation using a standard polynomial of degree 4.[9] The left panel shows the approximated function and the right panel the absolute difference to the true value function. We find that there are large approximation errors close to the singularities at $\lambda_2 = 0$ and $\lambda_2 = 1$.[10] In the third row we show the corresponding 4-degree approximation including the zero basis function $\Upsilon^0(\underline{\lambda}_2, s)$. We find that the approximation errors are several orders of magnitudes smaller. Also the errors are approximately equally distributed over the approximation interval, suggesting that the approximation adequately captures the properties of the singularities.

This concludes the description of the methodology for solving the heterogeneous agent model with recursive preferences. In the following section we apply the approach to solve the long-run risk model of Bansal and Yaron (2004) with heterogeneous agents.

[9] We use Chebychev nodes for the interpolation of the value function where the first node is fixed at 0 and the last node is fixed at 1.

[10] The errors for the value function of agent 1 close to $\lambda_2 = 1$ are significantly smaller in absolute values compared to the errors for the approximation of the value function of agent 2 close to $\lambda_2 = 0$. But the errors for agent 1 are by far largest close to the singularity $\lambda_2 = 1$.

Figure 2.1: Value Function Approximation in a Deterministic Economy



The figure shows the closed-form solution of the value function (2.38) as well as the consumption choices of the two agents (2.39) for $\underline{\lambda}_2 \in (0, 1)$. Aggregate consumption is constant at $C = 1$ and $\psi^1 = 1.5$ and $\psi^2 = 1.2$. The figure also shows 4-degree polynomial approximations of the value function and corresponding errors $V^h - \hat{V}^h$ with and without the zero basis function $\Upsilon^0(\underline{\lambda}_2, s)$.

2.3 Heterogeneous Agents and Long-Run Risks

In the model of Bansal and Yaron (2004) aggregate log consumption growth Δc_{t+1} is given by[11]

$$\begin{aligned} \Delta c_{t+1} &= \mu + x_t + \sigma\eta_{c,t+1} \\ x_{t+1} &= \rho x_t + \phi_x \sigma \eta_{x,t+1} \end{aligned} \tag{2.40}$$

where x_t denotes the long-run risk state and $\eta_{c,t+1}$ and $\eta_{c,t+1}$ are i.i.d. normal shocks. For the methodology described in this paper we need a process for the level of aggregate consumption. Therefore we transform the model from growth rates to levels. Let $c_t \equiv \log(C_t)$ denote the log of aggregate consumption. Model (2.40) can then equivalently be written as

$$\begin{aligned} c_{t+1} &= g_{t+1} + \sigma\eta_{c,t+1} \\ g_{t+1} &= \mu + g_t + x_t + \sigma\eta_{c,t} \\ x_{t+1} &= \rho x_t + \phi_x \sigma \eta_{x,t+1} \end{aligned} \tag{2.41}$$

where g_t denotes the log growth rate of the economy that is integrated of order one and hence is not stationary. To obtain a stationary formulation of the model we rewrite all equations in terms of detrended variables and introduce the new state $c_t^* \equiv c_t - g_t = \sigma\eta_{c,t}$. Appendix 2.A shows how the equilibrium equations (MC)-(Dλ) can be modified to account for growth. The modification does not only apply to the model of Bansal and Yaron (2004) but can be used for any model with log growth rate g_t as long as $\Delta g_{t+1} = g_{t+1} - g_t$ is stationary.

Bansal and Yaron (2004) assume that there is a representative investor with EZ preferences given by equation (2.25). We relax this assumption and apply the methodology presented in this paper to solve the long-run risk model (2.41) with $H = 2$ agents that differ with respect to their EIS ψ^h and their risk aversion γ^h. The exercise serves to demonstrate the solution approach. While a full qualitative and quantitative analysis of the influence of agent heterogeneity in the long-run risk model is beyond the scope of this paper, the results should give some intuition about the potential effects of agent heterogeneity under the assumption of recursive preferences and provide insights for further research in this area. In the following we first describe the technical specifications that we use for the projection approach. Thereafter we show several results for different utility parameter specifications.

2.3.1 Algorithmic Ingredients

The long-run risk model (2.41) with $H = 2$ agents has three states. The endogenous detrended Negishi weight $\underline{\lambda}_t^{2,*}$, the exogenous long-run growth rate x_t and the, also exogenous, detrended

[11] For simplicity we use the first model considered in Bansal and Yaron (2004) that only features long-run risks but no stochastic volatility.

aggregate log consumption level c_t^* that we obtain by transforming the original long-run risk model from growth rates to levels. For the value function approximation with the projection approach we need to choose certain collocation nodes. In this example we use Chebychev nodes for all three states.[12] For the two exogenous states we choose the minimum and maximum values by taking three standard deviations around the unconditional means of the processes. For the endogenous state $\underline{\lambda}_t^{h,*}$ we use zero for the first and one for the last node to cover the entire state space.

We provide the solver with additional information that we can formally derive for the limiting cases. For example we know that for $\underline{\lambda}_t^{2,*} = 1$ ($\underline{\lambda}_t^{2,*} = 0$) agent 2 (1) consumes everything, so it corresponds to the representative agent economy populated only by agent 2 (1). Hence, we require that the value function for these cases equals the value function for the corresponding representative agent economy. We also know that for $\underline{\lambda}_t^{2,*} = 0$ ($\underline{\lambda}_t^{2,*} = 1$) consumption of agent 2 (1) is 0 and hence the value function is zero as long as $\psi^h > 1$. In the case of $\psi^h < 1$ the value function becomes $-\infty$ which is an unpleasant property when polynomials are used for the function approximation. Therefore, instead of V^h, we approximate $\frac{1}{V^h}$ in the case of $\psi^h < 1$. By this approach we obtain very robust approximations for the value function even for $\psi^h < 1$ as the following examples show. We use a degree-10 polynomial for the $\underline{\lambda}_t^{h,*}$ dimension, a degree-4 polynomial for the x_t dimension and a degree-2 polynomial for the c_t^* dimension. As the shocks in the model are normally distributed, we compute the expectations over the exogenous states by Gauss-Hermite quadrature using 5 nodes for x_t and 3 nodes for c_t^*. For the following examples we use the parametrization from Bansal and Yaron (2004) given by $\mu = 0.0015, \rho = 0.979, \sigma = 0.0078, \phi_x = 0.044$ and $\delta = 0.998$. Also we assume that the first agent is the representative investor from Bansal and Yaron (2004) with a risk aversion of $\gamma^1 = 10$ and an EIS of $\psi^1 = 1.5$ and for the second agent we consider different utility parameter combinations to analyze the influence on model outcomes.

2.3.2 Results

We begin with a short exercise to demonstrate the accuracy of the solution method. In Table 2.1 we show absolute errors in the value function approximation (2.32) for different utility parameters of the two agents.[13] We find that for the low degree approximation excluding the zero basis function approximation errors are rather large with maximum errors as large as 0.0306. Including the zero basis functions significantly reduces approximation errors. We observe the same pattern for the high degree approximation where the errors are reduced by about one order of magnitude when including the zero basis function. For the two cases where

[12] See Chapter 4, Section 4.2.1.2 for a description of Chebychev nodes.

[13] To compute the errors, we set up a tensor grid for the three states at which we evaluate the value function approximation (2.32). We use 30 equally spaced nodes for the endogenous state and 15 nodes for each exogenous state. We report errors with and without the zero basis function (2.36). Also we consider two different polynomial degrees. The first approximation, denoted 'High Degree', uses a degree-10 polynomial for the $\underline{\lambda}^{2,*}$ dimension while the second approximation, denoted 'Low Degree', uses a degree-6 polynomial. Both approximations use a degree-4 polynomial for x and a degree-2 polynomial for c^*.

the second agent has an EIS smaller than one, and we approximate $\frac{1}{V^2}$ instead of the value function V^2 itself, approximation errors are slightly larger compared to the cases with an EIS larger than one. However, for all utility parameter combinations, approximation errors for the high degree approximation including the zero basis function are low with a maximum error of 4.41e-4 suggesting a high accuracy of our computational solution approach also for an EIS smaller than one.

Table 2.1: Errors Value Function

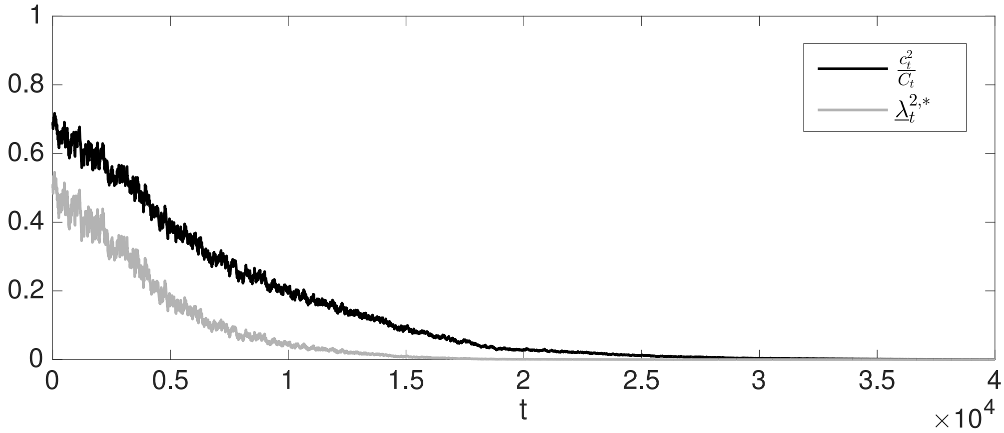
	High Degree		Low Degree	
	w ZBF	w/o ZBF	w ZBF	w/o ZBF
$\psi^1 = 1.5, \gamma^1 = 10$	5.89e-4	0.0030	0.0018	0.0047
$\psi^2 = 0.5, \gamma^2 = 10$	4.41e-4	0.0043	0.0060	0.0306
$\psi^1 = 1.5, \gamma^1 = 10$	6.41e-5	9.82e-5	2.95e-4	7.52e-4
$\psi^2 = 1.5, \gamma^2 = 8$	5.70e-5	8.91e-5	2.38e-4	6.77e-4
$\psi^1 = 1.5, \gamma^1 = 10$	7.84e-4	0.0037	0.0024	0.0060
$\psi^2 = 0.5, \gamma^2 = 8$	5.66e-4	0.0044	0.0064	0.0243

The table shows absolute errors in the value function approximation (2.32) *for different utility parameter combinations. The 'High Degree' approximation uses a degree-10 polynomial for the* $\underline{\lambda}^{2,*}$ *dimension and the 'Low Degree', uses a degree-6 polynomial. Both approximations use a degree-4 polynomial for* x *and a degree-2 polynomial for* c^**. Errors are shown with and without the zero basis function* (2.36).

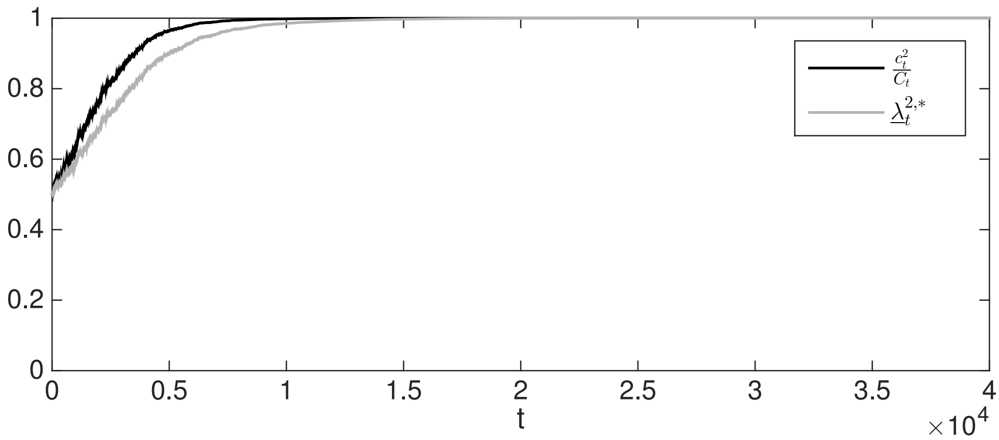
In the following we show how equilibrium outcomes are influenced by the agent heterogeneity. We begin by analyzing if and how multiple agents survive in the long-run risk model. For this, we show in Figure 2.2 the Negishi weight $\underline{\lambda}_t^{2,*}$ and the consumption share $\frac{c_t^2}{C_t}$ for 40,000 months of simulated data starting at $\underline{\lambda}_0^{2,*} = 0.5$. We consider three different utility specifications. In all three cases the first agent is the representative agent from Bansal and Yaron (2004) with a risk aversion of $\gamma^1 = 10$ and an EIS of $\psi^1 = 1.5$. In Panel (a) we assume that the second agent has the same risk aversion $\gamma^2 = 10$ but an EIS of $\psi^2 = 0.5$. We observe that the consumption share of agent 2 goes to zero in the long-run. The reason for this is that, while both agents are willing to take the same risk, agent 1 has a greater incentive to save as $\psi^1 > \psi^2$, which is why he accumulates all wealth in the long-run. In Panel (b) we assume that agent 2 has the same EIS as agent 1 $\psi^1 = \psi^2 = 1.5$, but he is less risk averse with $\gamma^2 = 8$. We find that in this case agent 2 dominates the economy as he is willing to take more risk, that in turn is rewarded with a higher return on average. Hence, in the long run only agent 2 survives and the consumption share of agent 1 goes to zero. In Panel (c) we show the case where agent 2 is less risk averse $\gamma^2 = 8$ but he also has a smaller EIS $\psi^2 = 0.5$ compared to agent 1. Or put differently, he has less incentive to safe, but is eager to take more risk. We observe that in this setup both agents survive and the long-run consumption share of agent 2 is between 0.6 and 0.8. So for recursive preferences, a difference in risk aversion can be offset by a difference in

Figure 2.2: Consumption Shares in the Long-Run Risk Model with Two Agents

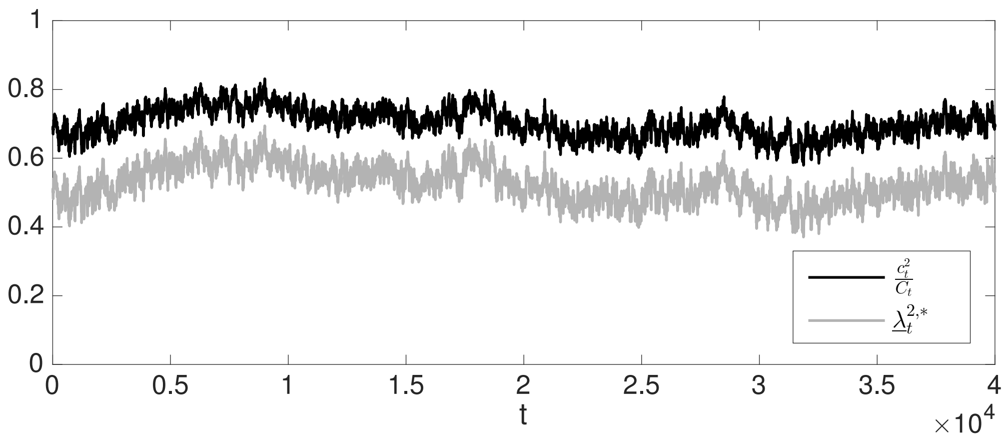
(a) $\psi^1 = 1.5, \gamma^1 = 10, \psi^2 = 0.5, \gamma^2 = 10$



(b) $\psi^1 = 1.5, \gamma^1 = 10, \psi^2 = 1.5, \gamma^2 = 8$



(c) $\psi^1 = 1.5, \gamma^1 = 10, \psi^2 = 0.5, \gamma^2 = 8$



The figure shows Negishi weights and the consumption shares for the long-run risk model of Bansal and Yaron (2004) with two agents that differ with respect to their preference parameters. The variables are shown for $40,000$ months of simulated data starting at $\underline{\lambda}_0^{2,*} = 0.5$.

the EIS to ensure survival of multiple agents.[14] This finding is fundamentally different from the results for standard time separable preferences, where a difference in risk aversion certainly leads to extinction of the more risk averse agents in the long-run (see e.g. Bhamra and Uppal (2014)).[15]

To better understand how and why multiple agents survive, we compute the wealth-consumption ratios of the individual agents. The wealth-consumption ratio for EZ preferences is given by[16]

$$\frac{W_t^h}{c_t^h} = \frac{1}{1-\delta^h}\left(\frac{V_t^{h,EZ}}{c_t^h}\right)^{1-\frac{1}{\psi^h}}. \tag{2.42}$$

For the reformulated value function V_t^h (see equation (2.25)) we obtain

$$\begin{aligned}\frac{W_t^h}{c_t^h} &= \frac{1-\frac{1}{\psi^h}}{1-\delta^h}\frac{V_t^h}{\left(c_t^h\right)^{1-\frac{1}{\psi^h}}} \\ &= \frac{1-\frac{1}{\psi^h}}{1-\delta^h}\frac{V_t^{h,*}}{\left(c_t^{h,*}\right)^{1-\frac{1}{\psi^h}}}.\end{aligned} \tag{2.43}$$

From the individual wealth-consumption ratios we can also compute the aggregate wealth-consumption ratio in the economy given by

$$\frac{W_t}{C_t} = \sum_{h=1}^{H}\frac{W_t^h}{c_t^h}\cdot\frac{c_t^h}{C_t}. \tag{2.44}$$

In Figure 2.3 we show the consumption shares and the wealth-consumption ratios for $\gamma^1 = 10, \psi^1 = 1.5, \gamma^2 = 8$ and $\psi^2 = 0.5$ where both agents survive in the long-run. We initially equip either of the agents with only a very small consumption share to analyze how they escape distinction. For this we simulate the model starting at $\underline{\lambda}_0^{2,*} = 0.0001$ and $\underline{\lambda}_0^{2,*} = 0.99$. We 'turn off' the exogenous shocks ($\eta_{c,t} = \eta_{x,t} = 0 \forall t$) for a better visualization.[17] Panel (a) shows the Negishi weight $\underline{\lambda}_t^{2,*}$ and the consumption share $\frac{c_t^2}{C_t}$ for $\underline{\lambda}_0^{2,*} = 0.0001$. We observe that the consumption share of agent 2 remains quite small for a while, but eventually increases

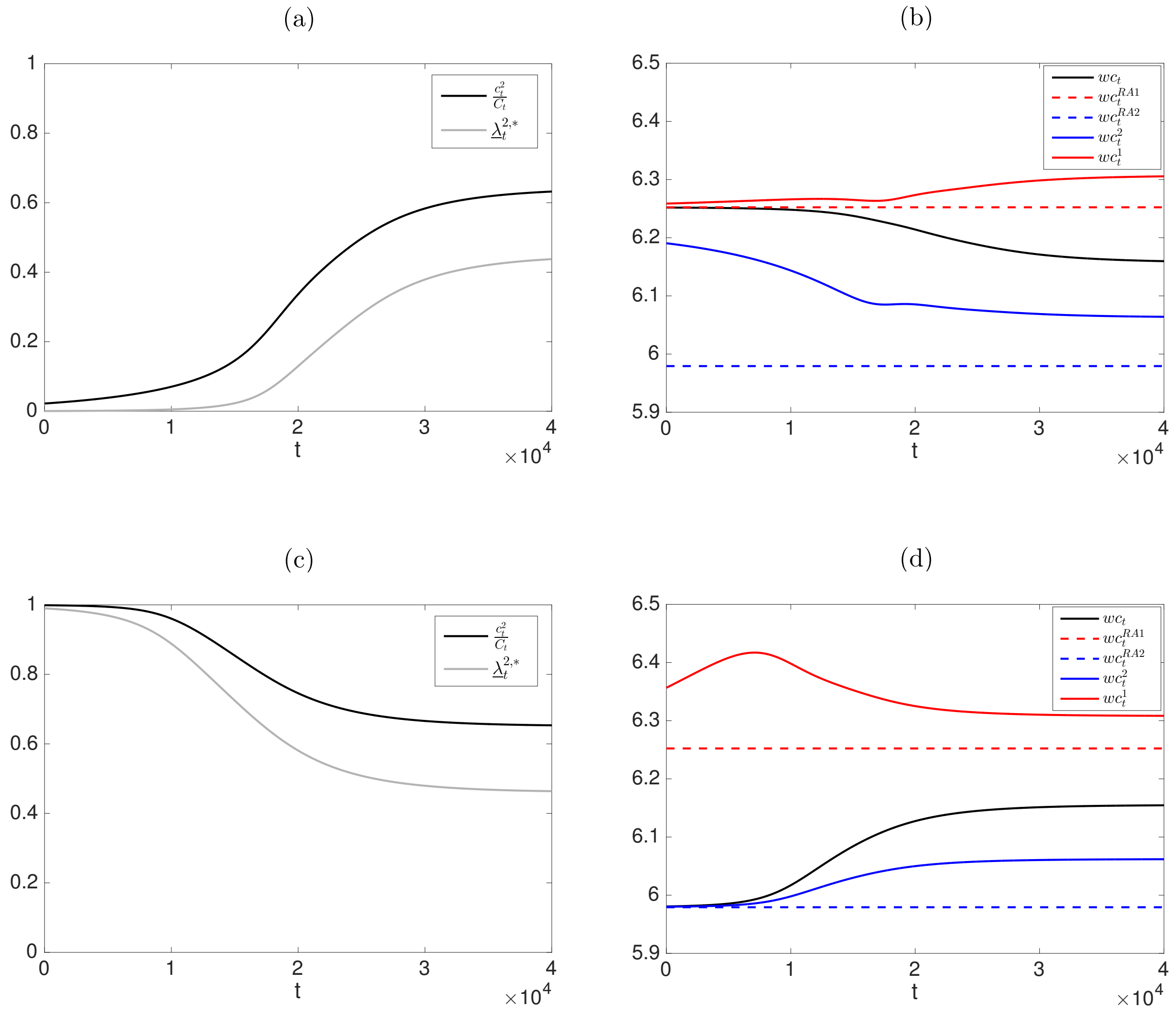
[14] These results are in line with Branger, Dumitrescu, Ivanova, and Schlag (2011) who provide utility parameter regions that ensure survival for the long-run risk model with two agents in a continuous time setup.

[15] In the economy with time separable preference the difference in risk aversion can, in theory, be offset by a difference in the subjective time preference ensuring survival of both agents. However, this is not true for a region of parameters but only for certain knife-edge combinations of the preference parameters.

[16] See the lecture notes of François Gourio on recursive utility, page 6, available at `https://sites.google.com/site/fgourio/teaching-notes`.

[17] We obtain the same qualitative patterns when the shocks are included but with a lot more 'noise' around the paths that makes it more difficult to extract meaningful insights from the graphs.

Figure 2.3: Survival Dynamics in the Long-Run Risk Model with Two Agents



The figure shows Negishi weights, consumption shares and log wealth-consumption ratios for the long-run risk model of Bansal and Yaron (2004) with two agents and $\gamma^1 = 10, \psi^1 = 1.5, \gamma^2 = 8$ and $\psi^2 = 0.5$. The variables are shown for 40,000 months of simulated data and there are no exogenous shocks ($\eta_{c,t} = \eta_{x,t} = 0 \forall t$). Panels (a) and (b) show the results for an initial $\underline{\lambda}_0^{2,*} = 0.0001$ and panels (c) and (d) for $\underline{\lambda}_0^{2,*} = 0.99$.

sharply and converges to its steady state slightly above 0.6 (in line with Figure 2.2 (c)). Panel (b) shows the corresponding log wealth-consumption ratios wc_t^h of the two agents, the aggregate log wealth-consumption ratio wc_t as well as the two log wealth-consumption ratios wc_t^{RAh} of a representative agent economy populated by either agent 1 or 2.

As the consumption share of agent 1 is close to one initially, his wealth-consumption ratio almost equals the aggregate wealth-consumption ratio (and also the wealth-consumption ratio of the representative agent economy populated only by agent 1) as the influence of agent 2 is negligible. So how does agent 2 overcome extinction and eventually regain his long-run consumption share? Agent 2 is less risk averse compared to agent 1. He uses his initial wealth to invest in the risky asset with the higher expected return to increase his consumption in the long-run. So his wealth-consumption ratio decreases in the beginning to escape extinction. In contrast with $\underline{\lambda}_0^{2,*} = 0.99$, where agent 1 is close to extinction (panels (c) and (d)), the wealth-consumption ratio of agent 1 is increasing initially. Agent 1 has a larger EIS compared to agent 2 and therefore has a greater incentive to save. So agent 1 escapes extinction by saving more and increasing his wealth-consumption ratio in the beginning. Once he has reached a certain wealth level and the risk of extinction has vanished, his consumption increases. This in turn lowers his wealth-consumption ratio, and the economy eventually converges to the long-run steady state.

Finally, in Figure 2.4 we show the time series including shocks for 600 months of simulated data. We initialize the time series at the long-run mean level of $\underline{\lambda}_t^{2,*}$ given by $\underline{\lambda}_{LR}^{2,*} = 0.46$ to analyze the wealth-consumption ratios of the individual agents in the long-run equilibrium.[18] Additional to the Negishi weights, consumption shares and wealth-consumption ratios we also show the long-run risk state x_t. In Table 2.2 we show corresponding correlation coefficients and standard deviations for the variables.

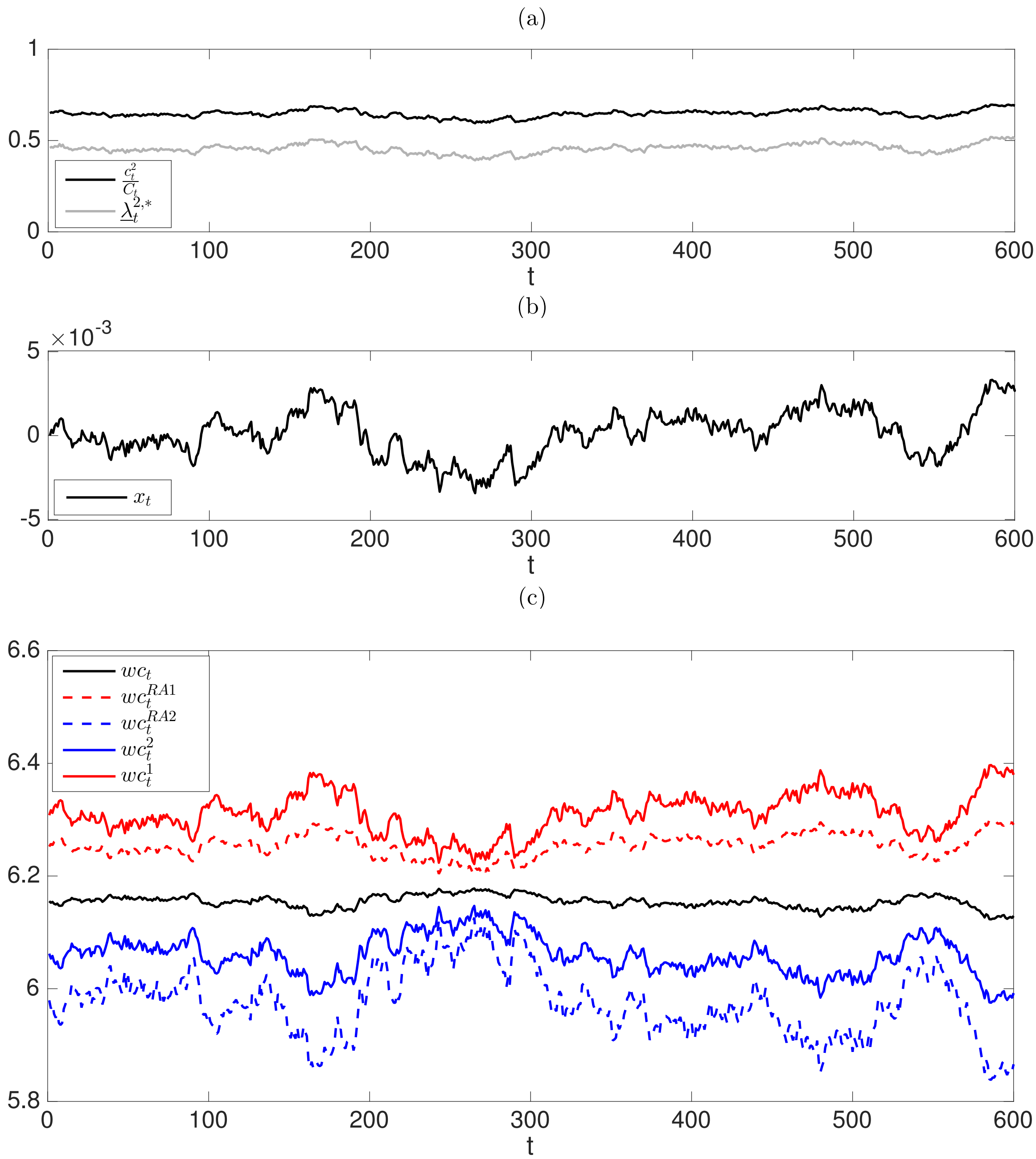
Table 2.2: Correlation Coefficients and Standard Deviations of Key Variables in the Long-Run Risk Model with Two Agents

	x_t	wc_t^1	wc_t^2	wc_t	wc_t^{RA1}	wc_t^{RA2}
x_t	0.0014	0.9997	-0.9998	-0.9890	1	-1
wc_t^1		0.0360	-1.0	-0.9923	0.9997	-0.9997
wc_t^2			0.0352	0.9920	-0.9998	0.9997
wc_t				0.0112	-0.9891	0.9888
wc_t^{RA1}					0.0201	-1
wc_t^{RA2}						0.0591

The table shows correlation coefficients (off diagonal elements) and standard deviations (diagonal elements) for the long-run risk model of Bansal and Yaron (2004) with two agent and $\gamma^1 = 10, \psi^1 = 1.5, \gamma^2 = 8$ and $\psi^2 = 0.5$. The values are computed using 600 months of simulated data starting at $\underline{\lambda}_0^{2,*} = 0.46$.

[18] We compute the long-run mean level of $\underline{\lambda}_t^{2,*}$ by simulating 1000 paths of 40,000 years of data each and computing the mean of $\underline{\lambda}_{40,000}^{2,*}$ over all paths.

Figure 2.4: Wealth-Consumption Ratios in the Long-Run Risk Model with Two Agents



The figure shows Negishi weights, consumption shares, the long-run risk state and log wealth-consumption ratios for the long-run risk model of Bansal and Yaron (2004) with two agents and $\gamma^1 = 10, \psi^1 = 1.5, \gamma^2 = 8$ and $\psi^2 = 0.5$. The variables are shown for 600 months of simulated data starting at $\underline{\lambda}_0^{2,*} = 0.46$.

We observe that the log wealth-consumption ratio of agent 1 (2) is positively (negatively) correlated with x_t. This result is in line with Bansal and Yaron (2004), who show that an EIS larger than 1 is required for a positive influence of the long-run risk state on the wealth-consumption ratio. As the two individual wealth-consumption ratios move in opposite directions and there is no large volatility in the consumption shares, the volatility of the aggregate wealth-consumption ratio is lower compared to the two individual wealth-consumption ratios as well as the wealth-consumption ratios of the representative agent economies. Due to its lower risk aversion and the willingness to invest more in the risky asset with the higher return, agent 2 has a consumption share of about 0.65 in the long-run. Consequently, the aggregate log wealth-consumption ratio wc_t is mainly driven by agent 2 as can be observed by the positive correlation with wc_t^2 and the negative correlation with wc_t^1.

This concludes the example to demonstrate the solution approach. We acknowledge that the results do not provide an extensive assessment of agent heterogeneity in the long-run risk model. But the example shows that there are interesting new features introduces by differences in agents with recursive preferences. Hence, further research needs to be conducted to fully understand the qualitative and quantitative implications of agent heterogeneity on equilibrium outcomes.

2.4 Conclusion and Outlook

We have presented a general methodology to solve asset pricing models with heterogeneous agents and recursive preferences. Recent findings from the asset pricing literature (e.g. Bansal and Yaron (2004)) highlight the importance of recursive preferences to explain long-standing asset pricing puzzles. So it is of great importance to better understand the implications of agent heterogeneity in this context. The solution method provided in this paper can be used for this task. Interesting questions that can be addressed with the approach are for example the analysis of survival dynamics and market selection, or the effects on the wealth distribution, the pricing kernel and financial markets outcomes. Also extending the framework to differences in beliefs among the agents is an interesting direction for future research.

2.A Including Growth in the Equilibrium Dynamics

In the following lines we show how the equilibrium equations (MC)-(Dλ) can be modified to account for growth. Let g_t denote the log growth rate of the economy and we require that $\Delta g_{t+1} = g_{t+1} - g_t$ is stationary. To obtain a stationary formulation of the equilibrium we rewrite all equations in terms of detrended variables. Let $c_t^* \equiv c_t - g_t$ denote detrended log aggregate consumption and $c_t^{h,*} \equiv \frac{c_t^h}{e^{g_t}}$ detrended individual consumption. The market clearing condition (MC) then becomes

$$\sum_{h=1}^{H} c_t^{h,*} = e^{c_t^*}.$$

To derive the optimality condition for the individual consumption decisions (CD) in terms of the detrended variables we consider the original specification of the equation with non-normalized λ_t^h:

$$\lambda_t^h(1-\delta^h)(c_t^h)^{-\frac{1}{\psi^h}} = \lambda_t^1(1-\delta^1)(c_t^1)^{-\frac{1}{\psi^1}}, \quad h \in \mathbb{H}^-.$$

Inserting $c_t^h = c_t^{h,*}e^{g_t}$ yields

$$\lambda_t^{h,*}(1-\delta^h)(c_t^{h,*})^{-\frac{1}{\psi^h}} = \lambda_t^{1,*}(1-\delta^1)(c_t^{1,*})^{-\frac{1}{\psi^1}}$$

where we define $\lambda_t^{h,*} \equiv \frac{\lambda_t^h}{e^{\frac{1}{\psi^h}g_t}}$. We normalize $\underline{\lambda}_t^{h,*} \equiv \frac{\lambda_t^{h,*}}{\lambda_t^{1,*}+\lambda_t^{h,*}}$ to obtain

$$\underline{\lambda}_t^{h,*}(1-\delta^h)(c_t^{h,*})^{-\frac{1}{\psi^h}} = (1-\underline{\lambda}_t^{h,*})(1-\delta^1)(c_t^{1,*})^{-\frac{1}{\psi^1}}.$$

So we end up with a new state variable $\underline{\lambda}_t^{h,*} \in (0,1)$, instead of $\underline{\lambda}_t^h$, where the following relationship can be derived by simple algebra:

$$\frac{1-\underline{\lambda}_t^{h,*}}{\underline{\lambda}_t^{h,*}} e^{(\frac{1}{\psi^1}-\frac{1}{\psi^h})g_t} = \frac{1-\underline{\lambda}_t^h}{\underline{\lambda}_t^h}. \tag{2.45}$$

Next we rewrite equations (VF) and (Dλ) in terms of the detrended variables. We define $V_t^{h,*} \equiv \frac{V_t^h}{e^{\left(1-\frac{1}{\psi^h}\right)g_t}}$. The value functions of the individual agents can then be expressed as

$$V_t^{h,*} = (1-\delta^h)\frac{(c_t^{h,*})^{1-\frac{1}{\psi^h}}}{1-\frac{1}{\psi^h}} + \delta^h R_t^h\left[e^{(1-\frac{1}{\psi^h})\Delta g_{t+1}}V_{t+1}^{h,*}\right]. \tag{2.46}$$

So we solve the model in terms of the detrended value function $V_t^{h,*}$.[19]

Finally, we derive the dynamics of $\underline{\lambda}_t^{h,*}$. We have

$$\Pi_{t+1}^h = \delta^h \left(\frac{V_{t+1}^h}{R_t^h \left[V_{t+1}^h \right]} \right)^{\frac{\frac{1}{\psi^h} - \gamma^h}{1 - \frac{1}{\psi^h}}} = \delta^h \left(\frac{V_{t+1}^{h,*}}{R_t^h \left[V_{t+1}^{h,*} \right]} \right)^{\frac{\frac{1}{\psi^h} - \gamma^h}{1 - \frac{1}{\psi^h}}}$$

so the expression for Π_{t+1}^h does not change. Rearranging equation (Dλ) gives

$$\frac{1 - \underline{\lambda}_{t+1}^h}{\underline{\lambda}_{t+1}^h} = \frac{\Pi_{t+1}^1}{\Pi_{t+1}^h} \frac{1 - \underline{\lambda}_t^h}{\underline{\lambda}_t^h}$$

and inserting (2.45) yields

$$\frac{1 - \underline{\lambda}_{t+1}^{h,*}}{\underline{\lambda}_{t+1}^{h,*}} e^{(\frac{1}{\psi^1} - \frac{1}{\psi^h})\Delta g_{t+1}} = \frac{\Pi_{t+1}^1}{\Pi_{t+1}^h} \frac{1 - \underline{\lambda}_t^{h,*}}{\underline{\lambda}_t^{h,*}}.$$

Or equivalently:

$$\underline{\lambda}_{t+1}^{h,*} = \frac{\Pi_{t+1}^h \underline{\lambda}_t^{h,*}}{(1 - \underline{\lambda}_t^{h,*}) \Pi_{t+1}^1 e^{(\frac{1}{\psi^h} - \frac{1}{\psi^1})\Delta g_{t+1}} + \underline{\lambda}_t^{h,*} \Pi_{t+1}^h}.$$

This concludes the transformation of the first-order conditions to obtain a stationary formulation of the equilibrium.

[19]Note that in the long-run risk model (2.41) Δg_{t+1} is given by $\Delta g_{t+1} = g_{t+1} - g_t = \mu + x_t + c_t^*$ so it only depends on time t information and hence we can move it out of the expression $R_t^h[\cdot]$.

Essay 3

Temporary Shocks and Asset Prices

Asset Prices with Temporary Shocks to Consumption[1]

Karl Schmedders
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
karl.schmedders@business.uzh.ch

Walter Pohl
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
walter.pohl@business.uzh.ch

Ole Wilms
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
ole.wilms@business.uzh.ch

December 2014

Abstract

Most standard asset pricing models assume that all shocks to consumption are permanent. We relax this assumption and allow also for temporary shocks. The implications of our model are dramatically different from those obtained in the prior literature. A canonical and parsimonious asset pricing model with CRRA preferences and temporary shocks can reproduce the equity premium, high return volatility and return predictability with a coefficient of relative risk aversion below ten. This finding suggests that temporary shocks can play an important role in explaining asset pricing puzzles.

Keywords: Asset prices; Equity premium; Unit root; Temporary shocks.

JEL Classification: G11, G12.

Note: This paper has been submitted to the Journal of Economic Dynamics and Control.

[1] We are indebted to David Andolfatto, George Constantinides, Nikola Gradojevic, Lars Hansen, Ken Judd, Martin Lettau, and Ilias Tsiakas for helpful discussions on the subject. We thank seminar audiences at Stanford, the University of Zurich, the Becker Friedman Institute at the University of Chicago, the 2014 Rimini Conference in Economics and Finance, and the 2014 World Finance Conference in Venice for comments. Karl Schmedders gratefully acknowledges financial support from the Swiss Finance Institute. Ole Wilms gratefully acknowledges financial support by Forschungskredit of the University of Zurich.

3.1 Introduction

It is a well-understood fact in the asset pricing literature that if shocks to consumption are permanent – so consumption is a random walk – then the behavior of U.S. stock prices are difficult to reconcile with a parsimonious model of agents with constant relative risk aversion (CRRA). Expected returns are too high, stock prices are too volatile, and future returns are too predictable to be generated by such a model given the low volatility of consumption growth in the data. Thus, the general thrust of asset pricing models has been towards more complex elements such as models with external habit or long-run risk, or towards disaster risk; see, among other papers, Campbell and Cochrane (1999), Wachter (2006), and Santos and Veronesi (2010); Bansal and Yaron (2004), Hansen, Heaton, and Li (2008), and Bansal and Shaliastovich (2013); Barro (2006), Rodriguez (2006), and Nakamura, Steinsson, Barro, and Ursúa (2013).

But what if there are temporary shocks to consumption, shocks whose impact diminishes over time? Then, as we show in this paper, the aforementioned conclusions neatly reverse themselves. Even for moderate levels of risk aversion in a canonical and parsimonious model, stock prices are volatile, expected returns are high, and future stock returns are partially predictable.

The question of whether shocks to the economy are temporary or permanent has led to a long and controversial discussion. Different studies (Nelson and Plosser (1982), DeJong, Nankervis, Savin, and Whiteman (1992)) have come down on both sides of the issue. Distinguishing between permanent and temporary-yet-persistent consumption shocks is very difficult given the data we have. Fortunately, our results are not driven by the *absence* of a permanent shock, but by the *presence* of temporary shocks. A model with a mixture of permanent and temporary shocks exhibits similar behavior to one with temporary shocks alone.

We amass several pieces of evidence towards the importance of temporary shocks for asset pricing. We consider a general model of consumption that experiences both permanent and temporary shocks. One special case of this model – where temporary shocks only last for one period – permits an exact analytical solution. We exploit this case to show how including a temporary shock can produce both a high equity premium and volatile returns with moderate levels of risk aversion.

We then calibrate a parsimonious model with a single shock to U.S. consumption and return data. The shock is very persistent, but not permanent. We choose three different empirical targets for our calibration exercise. In our base case, we use consumption data from 1889 to the present. Explaining the equity premium in post-war data is particularly challenging, so we also consider post-war data as a second target. Finally, an emerging literature (Savov (2011), Da and Yun (2011), Qiao (2013)) argues that mismeasurement in consumption has led to an artificially smooth consumption series, so we consider Savov's proxy consumption series as a third empirical target. In all three cases, we are able to produce a high equity premium and volatile returns with much lower levels of risk aversion than would be required if the shock were permanent. The post-war NIPA consumption data does indeed require a somewhat higher level of risk aversion than the long sample, but the long sample and Savov's proxy consumption series lead to very similar results. We also show that adding an additional permanent shock in these calibrations does not materially change the results.

Models with temporary consumption shocks are also able to generate many other time-series properties of asset prices with low levels of risk aversion. For example, in the temporary-shock model most variation in the price-dividend ratio is driven by changes in expected returns rather than changes in expected dividends. Temporary shocks are also sufficient to generate return predictability and meet the Hansen-Jagannathan bounds.

Several authors have previously considered the asset pricing implications of trend-stationary consumption. Tallarini (2000) considers mean reversion in consumption for Epstein-Zin utility with the special case where the elasticity of intertemporal substitution is one, where the model is exactly solvable. That paper finds that for high levels of risk aversion mean reversion in consumption actually lowers the equity premium. DeJong and Ripoll (2007) estimate a model with trend-stationary dividends and consumption, but use a log-linear approximation that leads to a much smaller estimate of the equity premium. Rodriguez (2006) finds that trend-stationary consumption helps explain the volatilities of returns, but needs large possible permanent shocks (disaster states) to match the empirical equity premium. In particular in his calibrated model the probability of a drop in consumption of more than 25% exceeds 17% and the possibility of rare but large shocks to consumption explains a large fraction of the equity premium. Nakamura, Steinsson, Barro, and Ursúa (2013) incorporate large permanent and temporary rare disasters to consumption. When they occur, the temporary disasters cause an average drop of 11% in consumption per year and occur for six subsequent years (that is, a 11% drop each year, not a 11% drop over 6 years). The standard deviation of the disaster shocks is also quite large, which generates the risk of even larger losses in consumption. Our model contains no disaster shocks – the permanent and temporary shocks are calibrated to match the U.S. experience. Bansal, Kiku, and Yaron (2010) also find a large equity premium, but in a model that contains both short-term and long-term risks, as well as stochastic volatility. Alvarez and Jermann (2005) find, in interesting contrast to the equity market, that long-term bonds data suggests that shocks to consumption are permanent.

The remainder of this paper is organized as follows. In Section 3.2 we present the basic model and provide analytical solutions for the special case of one-period temporary shocks. Section 3.3 provides a description of the data sets and the consumption specifications for the baseline version of the model. In Section 3.4 we report results on the asset pricing implications of the model and demonstrate their robustness. Finally, Section 3.5 concludes. In the Appendix, we derive the analytical results, describe the numerical solution method, and report additional results.

3.2 A Consumption-Based Asset Pricing Model

We briefly describe the particular version of the standard Lucas (1978) asset pricing model that we employ in this paper with both temporary and permanent shocks. We then consider a special case that permits closed-form solutions for asset prices. We use this case to illustrate the impact of temporary shocks on asset prices.

3.2.1 Consumption and Asset Prices

We consider a standard Lucas (1978) infinite-horizon representative-agent asset pricing model in discrete time, $t = 0, 1, ...$ There is a single perishable consumption good in each period. The agent's consumption in period t is denoted by C_t. The agent has expected utility

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t u(C_t) \right],$$

with CRRA Bernoulli utility

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$$

and discount factor β. The stochastic discount factor, M_t, to price assets in this model is the well-known expression

$$M_t = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}. \tag{3.1}$$

The logarithm of consumption, $c_t = \log(C_t)$, is the sum of two processes, g_t and x_t, which are the permanent and temporary components of consumption, respectively.

$$\begin{aligned} c_t &= g_t + x_t \\ x_t &= \rho_c x_{t-1} + \sigma_\epsilon \epsilon_t \\ g_t &= \bar{g} + g_{t-1} + \sigma_\nu \nu_t \\ \epsilon_t, \nu_t &\sim N(0,1) \text{ i.i.d.} \end{aligned} \tag{3.2}$$

with $\rho_c < 1$ and $\bar{g}$ denoting the long-term expected growth rate of consumption.

This specification encompasses several processes that have been used in the economic literature. For $\rho_c = 0$, $\sigma_\epsilon = 0$, consumption is a simple random walk with drift as considered in many papers, see, for example, Mehra (2006) or Tallarini (2000). For $\sigma_\nu = 0$, consumption is trend-stationary which is the second process analyzed in Tallarini (2000).

We contrast our model with other complex consumption models in the literature. The long-run risk literature, beginning with Bansal and Yaron (2004), considers highly-persistent shifts in the long-run mean of consumption growth. In our specification, the long-run mean of consumption growth is a constant, and the dynamics are driven by short-run deviations of consumption from its long-run trend (g_t). To illustrate this feature of the specification, consider the logarithmic growth rates, $\Delta c_t = c_t - c_{t-1}$. Then

$$\Delta c_t = \bar{g} + (\rho_c - 1)x_{t-1} + \sigma_\epsilon \epsilon_t + \sigma_\nu \nu_t,$$

which shows that the growth rate process $\{\Delta c_t\}$ is correlated with the process of prior deviations from trend, $\{x_{t-1}\}$.

Nakamura, Steinsson, Barro, and Ursúa (2013) also decompose growth into temporary and permanent components, but both components feature large disasters. Those disasters cause an average drop of 11% in consumption per year and occur for six subsequent years (that is, a 11% drop each year, not a 11% drop over 6 years). The standard deviation of the disaster shocks is also quite large, which generates the possibility of huge losses in consumption. As a result, Nakamura, Steinsson, Barro, and Ursúa (2013) overestimate the volatility of consumption growth by a factor of 1.66. (3.5% in the data compared to 5.8% implied by the model) in the pre-war period. For post-war consumption, the difference is even larger with a factor of 2.78 (1.8% in the data compared to 5.0% implied by the model).

The expected value, standard deviation and the first-order autocorrelation of consumption growth are as follows,[2]

$$\begin{aligned} \mathrm{E}(\Delta c_t) &= \bar{g} && (3.3) \\ \sigma(\Delta c_t) &= \sqrt{\sigma_\nu^2 + \frac{2}{1+\rho_c}\sigma_\epsilon^2} && (3.4) \\ \mathrm{AC1}(\Delta c_t) &= \left(\frac{\rho_c - 1}{\rho_c + 1}\sigma_\epsilon^2\right)(\mathrm{Var}(\Delta c_t))^{-1}. && (3.5) \end{aligned}$$

The objective of this paper is to analyze the asset pricing implications of the model with the consumption process (3.2). In the first, baseline, version of the model we assume that there is no labor income and that there is a risky asset ("Lucas tree") paying dividends equal to the aggregate consumption claim,

$$D_t = C_t,$$

each period. We consider the consequences of relaxing this assumption and including a separate dividend process in Section 3.4.4 below. In the analysis of the baseline model we frequently call the risky asset the aggregate consumption claim.

For a one-period bond that pays one unit of the consumption good, the pricing equation reads

$$P_t^f = \beta E_t \left\{ \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \right\}. \tag{3.6}$$

[2] The derivations of the unconditional moments can be found in Appendix 3.A.1. We use these analytical expressions in our analysis below to exactly match the moments of the underlying consumption process to the data.

The return on the aggregate consumption claim can be expressed in terms of the price-consumption ratio,

$$\frac{P_t}{C_t} = \beta E_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \left(\frac{P_{t+1}}{C_{t+1}} + 1 \right) \right\}. \tag{3.7}$$

The one-period returns of the two assets are then given by

$$R^f_{t+1} = \frac{1}{P^f_t}, \tag{3.8}$$

$$R_{t+1} = \frac{\left(\frac{P_{t+1}}{C_{t+1}} + 1\right)}{\frac{P_t}{C_t}} \times \frac{C_{t+1}}{C_t}. \tag{3.9}$$

To numerically compute solutions for asset prices and returns, we rewrite the model in terms of the stationary variable $x_t = c_t - g_t$ and the change in the log growth rate $g_{t+1} - g_t = \bar{g} + \sigma_\nu \nu_{t+1}$. The pricing equation (3.6) for the riskless asset and the price-consumption ratio (3.7) for the infinitely-lived risky asset can then be expressed by

$$P^f_t = \beta E_t \left\{ e^{-\gamma(g_{t+1} - g_t + x_{t+1} - x_t)} \right\} \tag{3.10}$$

$$\frac{P_t}{C_t} = \beta E_t \left\{ e^{(1-\gamma)(g_{t+1} - g_t + x_{t+1} - x_t)} \left(\frac{P_{t+1}}{C_{t+1}} + 1 \right) \right\}. \tag{3.11}$$

Our calibration results lead us to consider models where $\beta > 1$. Kocherlakota (1990) first showed that in an economy with a positive growth rate, $\bar{g} > 0$, and sufficiently risk-averse investors, $\gamma > 1$, the discount factor β can, in fact, exceed one; the agent's utility maximization problem remains well-defined and a finite solution to the pricing equations (3.10) and (3.11) exists. In Appendix 3.A.2, we extend this result and show that the agent's expected utility remains finite for $\gamma > 1$ if

$$\beta < e^{(\gamma-1)\bar{g}}. \tag{3.12}$$

Piazzesi, Schneider, and Tuzel (2007) also consider values of $\beta > 1$.

3.2.2 Analytical Results: Permanent versus Temporary Shocks

As is well-known, if consumption simply follows a random walk, the model can be solved in closed form, see e.g. Mehra (2006). We can also solve the model for the more general consumption process (3.2) with both a temporary and a permanent shock as long as we assume that $\rho_c = 0$. This version of the model simplifies to

$$\begin{aligned} c_t &= g_t + \sigma_\epsilon \epsilon_t \\ g_t &= \bar{g} + g_{t-1} + \sigma_\nu \nu_t \\ \epsilon_t, \nu_t &\sim N(0,1) \text{ i.i.d.} \end{aligned} \tag{3.13}$$

For $\sigma_\epsilon = 0$, the model collapses to a simple random walk for log consumption. For $\sigma_\nu = 0$, log consumption is trend-stationary with persistence $\rho_c = 0$. (We consider non-zero ρ_c in our calibration results in Section 3.4.)

The key asset pricing moments are given by the following theorem.

Theorem 3. *Consider the model with the consumption process* (3.13) *exhibiting a temporary shock, ϵ_t, and a permanent shock, ν_t. Then the first and second unconditional moments of the risk-free rate and the expected return of the aggregate consumption claim are as follows,*

$$E\left(R_t^f\right) = \frac{1}{\beta} e^{\gamma\bar{g} - \frac{1}{2}\gamma^2\sigma_\nu^2} \tag{3.14}$$

$$E\left((R_t^f)^2\right) = \frac{1}{\beta^2} e^{2\gamma\bar{g} - \gamma^2\sigma_\nu^2 + \gamma^2\sigma_\epsilon^2} \tag{3.15}$$

$$E\left(R_{t+1}\right) = e^{\frac{1}{2}\sigma_\nu^2 + \gamma^2\sigma_\epsilon^2 + \bar{g}} + \frac{1 - \beta e^{(1-\gamma)\bar{g} + \frac{1}{2}(1-\gamma)^2\sigma_\nu^2}}{\beta} e^{\gamma\bar{g} + \gamma\sigma_\nu^2 - \frac{1}{2}\gamma^2\sigma_\nu^2 + \gamma\sigma_\epsilon^2} \tag{3.16}$$

$$\begin{aligned} E\left(R_{t+1}^2\right) &= e^{4\gamma^2\sigma_\epsilon^2 + 2\bar{g} + 2\sigma_\nu^2} + \frac{1}{X^2} e^{2\gamma^2\sigma_\epsilon^2 + 2\sigma_\epsilon^2 + 2\bar{g} + 2\sigma_\nu^2} \\ &+ \frac{2}{X} e^{2.5\gamma^2\sigma_\epsilon^2 + \gamma\sigma_\epsilon^2 + 0.5\sigma_\epsilon^2 + 2\bar{g} + 2\sigma_\nu^2}, \end{aligned} \tag{3.17}$$

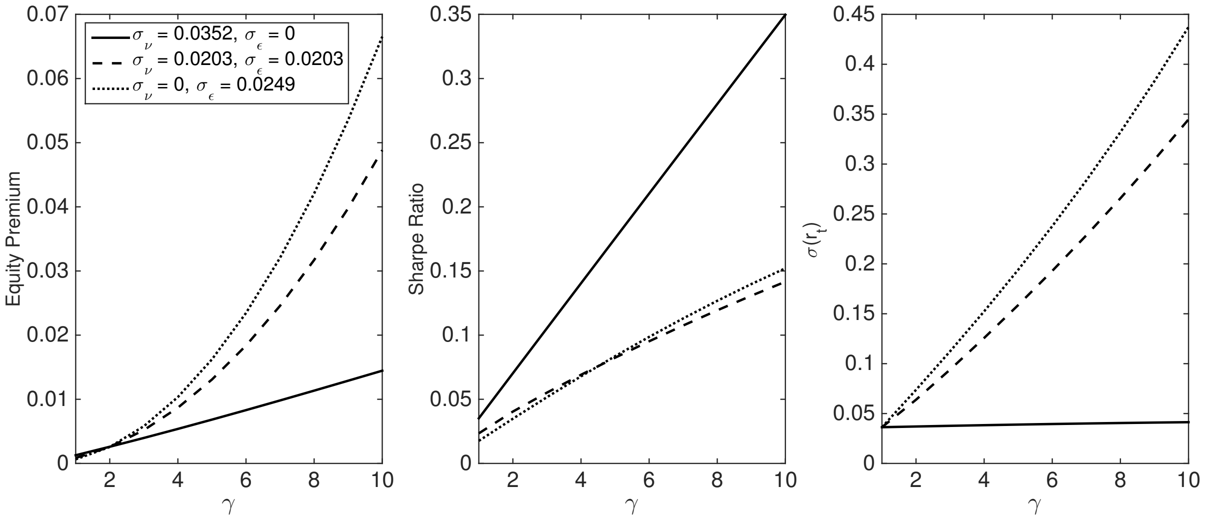
with the constant

$$X = \frac{\beta e^{(1-\gamma)\bar{g} + \frac{1}{2}(1-\gamma)^2(\sigma_\epsilon^2 + \sigma_\nu^2)}}{1 - \beta e^{(1-\gamma)\bar{g} + \frac{1}{2}(1-\gamma)^2\sigma_\nu^2}}.$$

Proof. See Appendix 3.A.3. □

Since the analytical expressions in Theorem 3 are rather complex, we illustrate them for a particular set of parameter values. We fix the expected growth rate at $\mathrm{E}(\Delta c_t) = 0.020$ and the volatility of consumption growth at $\sigma(\Delta c_t) = 0.0352$. (We derived these values from one of our three data sets, see Section 3.3 for a description of the data sets and Table 3.1 for the parameter estimates.) Figure 3.1 shows the equity premium, Sharpe ratio and volatility of the return on the aggregate consumption claim for different degrees of risk aversion γ. In all three graphs in Figure 3.1, the dotted line shows the case of $\sigma_\nu = 0$ when all consumption volatility comes from the temporary shock ϵ_t. The solid line shows the case of $\sigma_\epsilon = 0$ when all consumption volatility comes from the permanent shock ν_t. The dashed line shows the intermediate case $\sigma_\nu = \sigma_\epsilon$. The temporary shock generates a much higher equity premium and stock volatility than the permanent shock, particularly for higher levels of risk aversion. On the contrary, the Sharpe ratio is larger for the permanent shock than for the temporary shock. The large Sharpe ratio for the permanent shock, however, is not a consequence of a high excess return but instead of a very low return volatility. For a permanent shock the standard deviation of returns on the aggregate consumption claim is below 5% for all $\gamma \leq 10$, so even a small equity premium of 1.45% has a Sharpe ratio of 0.35. In contrast, for $\gamma = 10$ the temporary shock leads to a premium of 6.65% but a volatility of more than 40%.

Figure 3.1: Equity Premium, Sharpe ratio, and Return Volatility of the Consumption Claim



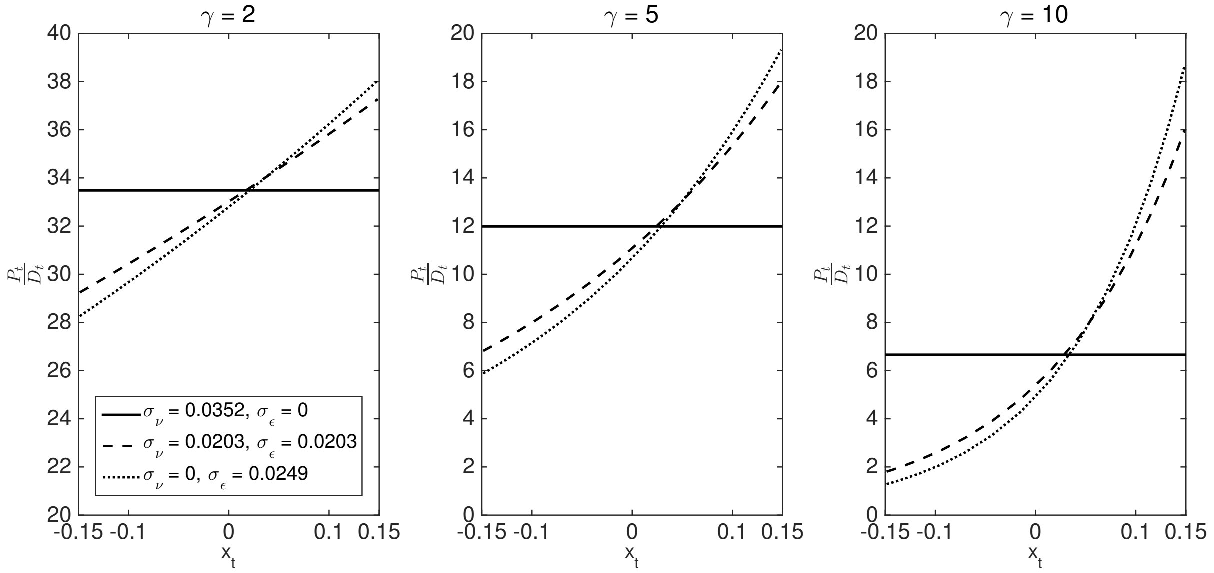
The graphs show the equity premium, Sharpe ratio and volatility of the return of the aggregate consumption claim for different degrees of risk aversion γ. The average growth rate is $\mathrm{E}(\Delta c_t) = 0.020$ and the volatility of consumption growth is $\sigma(\Delta c_t) = 0.0352$. We consider three sets of results. In the first we only have the persistent shock ν_t with $\sigma_\epsilon = 0$ (solid line). In the second we assume $\sigma_\epsilon = \sigma_\nu$ (dashed line) and in the third we have the case with only temporary shocks given by $\sigma_\nu = 0$ (dotted line). For all cases $\rho_c = 0$ and $\beta = 0.99$.

Figure 3.2 illustrates the price-dividend ratio as a function of the temporary shock. As risk aversion increases, the response to the shock becomes both larger and increasingly nonlinear. This demonstrates how a temporary shock is sufficient to generate interesting dynamics. The model with only a temporary shock and both temporary and permanent shocks show similar dynamic effects, while the model with only a permanent shock, of course, has no dynamics in the price-dividend ratio whatsoever.

This completes our initial analysis of the asset pricing implications of our economic model. Obviously, the special case of $\rho_c = 0$ does not reflect a property of actual market data but instead only serves as a benchmark to obtain a first impression of the different effects of temporary and permanent consumption shocks on asset prices. We find that temporary shocks produce significantly larger risk premia than permanent shocks and also increase the return volatilities of the assets. These properties come at the cost of much lower Sharpe ratios for temporary shocks than for permanent shocks.

Before we discuss more general results in Section 3.4 below, we first describe the properties of market data that are relevant for a proper specification of the consumption process.

Figure 3.2: Price-Dividend Ratio as a Function of the State x_t



The graph shows the price-dividend ratio $\frac{P_t}{D_t}$ as a function of the state x_t for three different degrees of risk aversion $\gamma = [2, 5, 10]$. The volatility of consumption growth is fixed at $\sigma(\Delta c_t) = 0.0352$ and $\mathrm{E}(\Delta c_t) = 0.020$. We consider three sets of results. In the first we only have the persistent shock ν_t with $\sigma_\epsilon = 0$ (solid line). In the second we assume $\sigma_\epsilon = \sigma_\nu$ (dashed line) and in the third we have the case with only temporary shocks given by $\sigma_\nu = 0$ (dotted line). For all cases $\rho_c = 0$ and $\beta = 0.99$.

3.3 Data and Summary Statistics

We describe our data sources and report summary statistics. Then we provide results from unit root tests on the data series for consumption, dividends, and asset prices, respectively. Finally we report parameter estimates for a trend-stationary consumption process and for a random walk specification.

3.3.1 Consumption, Dividends, and Return Series

We use U.S. consumption data to calibrate the underlying consumption process (3.2). Parameters are chosen so that the resulting moments (3.3)–(3.5) match those of observed market data. We consider three different aggregate consumption series to examine the consequences for our results.

The first series we consider is the annual consumption data series constructed by Robert J. Shiller.[3] Consumption is aggregate per-capita real personal consumption from the National Income and Product Accounts. (Prior to 1929 this data series is not available, so Shiller uses estimates from Kendrick (1961).) We refer to this sample as the "Long Sample".

[3]The dataset can be downloaded from `http://www.econ.yale.edu/~shiller/data.htm` (last accessed April 28, 2014).

The behavior of stock prices is particularly hard to explain in the post-World War II period (Grossman and Shiller (1981)). Therefore, we use as a second sample the Shiller consumption data from 1947–2009. Table 3.1 shows the mean, standard deviation, autocorrelation, and the correlation with the market portfolio of the consumption growth in the two time series.

Table 3.1: Empirical Moments of Consumption

	Long Sample	**Post-War**	**Garbage**
$E(\Delta c_t)$	0.0200	0.0213	0.0142
	(0.0032)	(0.0023)	(0.0041)
$\sigma(\Delta c_t)$	0.0352	0.0180	0.0286
	(0.0028)	(0.0014)	(0.0037)
$AC1(\Delta c_t)$	-0.0640	0.2466	-0.1438
	(0.1224)	(0.1286)	(0.1747)
Corr. R_m	0.5613	0.5454	0.6016
	(0.0631)	(0.0887)	(0.1110)

The table shows the mean $E(\Delta c_t)$*, standard deviation* $\sigma(\Delta c_t)$*, autocorrelation* $AC1(\Delta c_t)$ *and correlation with the market portfolio* Corr. R_m *of the different growth series. Bootstrapped standard errors from* 10^6 *simulations are provided in parentheses. The long sample consists of all real consumption from 1889–2009 and the post-war series from 1947–2009. The garbage data is available from 1960–2006.*

The moments of the two consumption series are in line with the values reported in Campbell and Cochrane (1999) or Guvenen (2009). The volatility in the post-war consumption series is significantly lower compared to the long sample. This property of the data is one of the reasons why it is so difficult to explain the large difference in equity and risk free returns in the post-war sample, see Grossman and Shiller (1981). Another source of difficulty is the positive autocorrelation in consumption growth rates.

An emerging literature (Savov (2011), Da and Yun (2011), Qiao (2013)) considers the consequences of mismeasurement in NIPA consumption. Triplett (1997) provides a critical look at how consumption is actually computed. Savov (2011) argues that the consumption estimates in the National Income and Product Accounts are artificially smooth. A smoothed series will have lower volatility than the true series, and the smoothing will introduce artificial positive autocorrelation, both of which make stock price dynamics harder to explain. Savov proposes municipal solid waste data collected by the U.S. Environmental Protection Agency as an alternative proxy for consumption. The logic is that consumption will generate waste, so waste should be highly correlated with actual consumption. In response to Savov's arguments, we also analyze the asset pricing implications of our model using the time series of garbage growth ranging from 1960–2006 as in Savov (2011).[4] We refer to this data series as "Garbage." The rightmost column in Table 3.1 reports the summary measures for this time series.

Risky asset prices are again taken from the Shiller website. Starting from 1926, risky asset pries are given by the January level of the S&P 500 (or its predecessor indices), deflated by

[4]We thank Alexi Savov for making his data available to us.

the January consumer price index (CPI-U). Dividends, D_t, are measured by the total amount of S&P 500 dividends in a year, deflated by the average CPI for that year. (Shiller again uses alternative sources to extend the data back to 1889. Stock data comes from Cowles and Associates (1939). The CPI-U series only extends back to 1913, so prior to that date Shiller uses the price index from Warren and Pearson (1935).) Table 3.2 reports the empirical moments of the dividends series for the respective time frames of our three data sets. Similarly, Table 3.3 reports the mean and standard deviation of the market return and the risk free rate (in percent) as well as the log price dividend ratio for the three different data sets.

Table 3.2: Empirical Moments of Dividends

	Long Sample	**Post-War**	**Garbage**
$\mathrm{E}(\Delta d_t)$	0.0106	0.0173	0.0144
	(0.0105)	(0.0081)	(0.0061)
$\sigma(\Delta d_t)$	0.1160	0.0647	0.0421
	(0.0131)	(0.0112)	(0.0034)
$\mathrm{AC1}(\Delta d_t)$	0.1379	0.4470	0.6836
	(0.1127)	(0.1677)	(0.0855)

The table shows the mean $\mathrm{E}(\Delta d_t)$*, standard deviation* $\sigma(\Delta d_t)$ *and autocorrelation* $\mathrm{AC1}(\Delta d_t)$ *of the dividend growth for the three different datasets. Bootstrapped standard errors from* 10^6 *simulations are provided in parentheses. The long sample extends from 1889–2009 and the post-war series from 1947–2009. For the garbage series, the moments are for the period of 1960–2006.*

Table 3.3: Empirical Moments of Financial Market Data

	$E(R_t)$	$\sigma(R_t)$	$E(R_t^f)$	$\sigma(R_t^f)$	$E(p_t - d_t)$	$\sigma(p_t - d_t)$
Long Sample	7.60	18.73	1.97	5.80	3.22	0.40
Post-War	7.92	16.60	1.84	2.65	3.42	0.44
Garbage	7.09	15.18	2.21	2.60	3.51	0.40

The table shows the mean and standard deviation of the market return and the risk free rate (in percent) as well as the log price dividend ratio for the three different data sets. The long sample extends from 1889–2009 and the post-war series from 1947–2009. For the garbage series, the moments are for the period of 1960–2006.

3.3.2 Unit Root Statistics

Before we can analyze the asset pricing model, we need to specify the parameters of the consumption process (3.2). This task forces us to confront the issue whether the consumption process has a unit root and to take a stand on the influence of temporary and permanent shocks on aggregate consumption.

The question as to whether shocks to the economy are temporary or permanent has led to a long and controversial discussion, ever since Nelson and Plosser (1982) first provided evidence that most macroeconomic time series have a unit root. DeJong and Whiteman (1991a), DeJong and Whiteman (1991b), DeJong and Whiteman (1991c), Kwiatkowski, Phillips, Schmidt, and Shin (1992) and DeJong, Nankervis, Savin, and Whiteman (1992) argue that for most macroeconomic time series the trend-stationarity hypothesis is much more likely than the unit root alternative. Perron (1989) and Andreou and Spanos (2003) present evidence that most macroeconomic time series are best represented by stationary fluctuations around a trend, with certain structural breaks, e.g., the 1929 crash or the 1973 oil price shock which both had persistent effects. Several authors (Christiano and Eichenbaum (1990), Cochrane (1991), Rudebusch (1993), and Diebold and Senhadji (1996)) observe that the presence or size of the persistent component is difficult to tease out with the data we have.

Clearly, this discussion in the literature is of great importance for the exact specification of the consumption process (3.2) in our model. Therefore, we conduct three common unit root tests for the time series of consumption, dividends, and asset prices for each of our three data sets. We employ the augmented Dickey-Fuller (ADF) (see Dickey and Fuller (1979)) and the Phillips and Perron (1988) (PP) test with the null hypothesis of a random walk with a constant and a drift, as well as the Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) test with the null hypothesis of trend-stationarity. Table 3.4 provides test statistics and critical values for the three tests.

Table 3.4: Test Statistics and Critical Values of the Unit Root Tests

		ADF-Test	PP-Test	KPSS-Test
Long Sample	c_t	-2.32	-2.58	0.20
	d_t	-4.15	-3.56	0.09
	p_t	-2.73	-2.59	0.15
Post-War	c_t	-2.38	-2.24	0.09
	d_t	-4.20	-3.62	0.09
	p_t	-1.95	-1.76	0.09
Garbage	c_t	-1.81	-1.85	0.13
	d_t	-2.34	-1.29	0.13
	p_t	-1.35	-1.31	0.14

	Critical Values								
	ADF 1%	ADF 5%	ADF 10%	PP 1%	PP 5%	PP 10%	KPSS 1%	KPSS 5%	KPSS 10%
	-3.99	-3.43	-3.13	-4.04	-3.45	-3.15	0.216	0.146	0.119

Test statistics for log consumption, dividends and prices of the ADF-Test with a constant and a trend using one lag order, the PP-Test where the truncation lag parameter is set to $trunc(4(T/100)^{0.25})$ *and the KPSS-Test where the truncation lag parameter is set to* $trunc(\frac{10}{14}\sqrt{T})$ *with* T *being the sample size. In the lower panel 1%, 5% and 10% critical values are provided.*

We find strong empirical evidence for trend-stationarity in the dividend series for the long sample and the post-war period, while it is not that obvious for consumption and prices. The null hypothesis of trend-stationarity in the KPSS test cannot be rejected for any of the time series at the 1% significance level. So neither the hypothesis of a unit-root nor the hypothesis of trend-stationarity can be ruled out by the tests. Therefore we analyze the asset pricing implications of the trend-stationary model against the unit-root hypothesis (random walk model).

Side Note. To emphasize the point that the random walk and trend-stationary model are almost indistinguishable by the tests, we run the unit root tests on simulated data. We simulate n observations of c_t where n is the length of the corresponding dataset (121 for the long sample, 63 for the post-war sample and 47 for the garbage sample). This is done 10,000 times. We report the median of the corresponding test statistics of the three tests as well as the standard deviations in parentheses. This is done for the case where consumption is trend-stationary and for the case where consumption is a random walk. Table 3.5 shows, that even in the simulated data, it is hard to distinguish the trend-stationary model from the random walk model. Looking at the results for the trend-stationary model ($\sigma_\nu = 0$) we find that both, the ADF-Test and the PP-Test, do not reject the null hypothesis of a random walk. We observe the same finding for the random walk model ($\sigma_\epsilon = 0$) and a model where the permanent shock accounts for 40% of the total volatility in consumption growth, $\frac{\sigma_\nu}{\sigma(\Delta c_t)} = 0.4$. (This third model is of interest to us below.) For all three models, we also cannot reject the null hypothesis of trend-stationarity in the KPSS test at a 5% significance level. In addition we note that the standard deviations of the test statistics are quite large, which suggests that the sample is too small to dismiss either one of the two alternatives. These results are in line with previous Monte Carlo studies by Schwert (2002).

3.3.3 Parameter Estimates for Consumption Processes

The critical input in the asset pricing model is the aggregate consumption process (3.2). In light of the results of the three unit root tests, we fit both a trend-stationary process and a random walk to the consumption time series for all three data sets. (In our robustness analysis below, we also consider models that contain both permanent and temporary shocks.)

3.3.3.1 Trend Stationary Consumption

For $\rho_c < 1$ the general (logarithmic) consumption specification (3.2) is trend-stationary when $\sigma_\nu = 0$ and so,

$$
\begin{aligned}
c_t &= g_t + x_t \\
x_t &= \rho_c x_{t-1} + \sigma_\epsilon \epsilon_t \\
g_t &= \bar{g} + g_{t-1} \\
\epsilon_t &\sim N(0,1) \text{ i.i.d.}
\end{aligned}
\tag{3.18}
$$

Table 3.5: Test Statistics and Critical Values of the Unit Root Tests for Simulated Consumption Data

		ADF-Test			**PP-Test**			**KPSS-Test**		
Long S.	$\sigma_\nu = 0$	-2.89 (0.65)			-3.02 (0.65)			0.11 (0.043)		
	$\sigma_\epsilon = 0$	-2.18 (0.79)			-2.24 (0.79)			0.14 (0.054)		
	$\frac{\sigma_\nu}{\sigma(\Delta c_t)} = 0.4$	-2.70 (0.68)			-2.80 (0.69)			0.12 (0.048)		
P.-War	$\sigma_\nu = 0$	-2.40 (0.72)			-2.51 (0.70)			0.11 (0.030)		
	$\sigma_\epsilon = 0$	-2.18 (0.80)			-2.24 (0.80)			0.12 (0.034)		
	$\frac{\sigma_\nu}{\sigma(\Delta c_t)} = 0.4$	-2.35 (0.74)			-2.45 (0.72)			0.11 (0.032)		
Garb.	$\sigma_\nu = 0$	-2.85 (0.71)			-3.04 (0.66)			0.10 (0.023)		
	$\sigma_\epsilon = 0$	-2.18 (0.81)			-2.26 (0.80)			0.12 (0.028)		
	$\frac{\sigma_\nu}{\sigma(\Delta c_t)} = 0.4$	-2.66 (0.75)			-2.83 (0.70)			0.11 (0.025)		
		Critical Values								
		1%	5%	10%	1%	5%	10%	1%	5%	10%
		-3.99	-3.43	-3.13	-4.04	-3.45	-3.15	0.216	0.146	0.119

The table shows statistics for the same tests as in Table 3.4 but for simulated consumption data. For this purpose we simulate n observations of c_t where n is the length of the corresponding dataset (121 for the long sample, 63 for the post-war sample and 47 for the garbage sample). We perform 10,000 such simulations. We report the median of the corresponding test statistics of the three tests as well as the standard deviations in brackets. This is done for the case where consumption is trend-stationary, $\sigma_\nu = 0$ (parameter estimates for the consumption process are taken from Table 3.6), for the case where consumption is a random walk, $\sigma_\epsilon = 0$ (parameter estimates are taken from Table 3.1 with $\sigma_\nu = \sigma(\Delta c_t)$) and for the case where the permanent shocks account for 40% of the total volatility in consumption growth, $\frac{\sigma_\nu}{\sigma(\Delta c_t)} = 0.4$.

This log consumption process is composed of a deterministic linear trend with AR(1) deviations. The smaller the coefficient ρ_c, the faster the process reverses to its linear trend. In the extreme case $\rho_c = 0$, the consumption process becomes a linear trend with white Gaussian noise. The larger the autocorrelation coefficient ρ_c, the more persistent is a shock ϵ_t to consumption. In the extreme case $\rho_c = 1$, the consumption process ceases to be trend-stationary and instead has a unit root and becomes a random walk with drift.

We estimate the three parameters $\bar{g}$, σ_ϵ, and ρ_c in the trend-stationary ("TS") consumption process (3.18) for each of our three data sets (long sample, post-war, garbage). For this purpose, we first remove from each consumption time series the linear trend $\bar{g} = \mathrm{E}(\Delta c_t)$ to obtain the de-trended time series $x_t = c_t - g_t$. Then the estimate for the coefficient ρ_c is simply the correlation coefficient of the de-trended levels and its one-period lag and σ_ϵ is the standard deviation of the residuals $x_t - \rho_c x_{t-1}$. Table 3.6 shows the point estimates as well as bootstrapped standard errors for the parameters for the different time series.

Recall from Table 3.1 that the volatility of consumption growth $\sigma(\Delta c)$ in the post-war series (1947–2009) and the garbage series (1960–2006) is much lower than in the long sample (1889–2009). As a result, the estimates for the standard deviation σ_ϵ is considerably larger for the long sample than for the post-war period and for the garbage data series. The estimates for the autocorrelation parameter ρ_c indicate that shocks to consumption are much less persistent for consumption based on garbage data than for the other two data series. In other words, the garbage data series exhibits the fastest reversion to long-run trend.

Table 3.6: Parameter Estimates for the Trend Stationary (TS) Consumption Process

	$\bar{g}$	σ_ϵ	ρ_c
Long Sample	0.0200	0.0343	0.9100
	(0.0032)	(0.0024)	(0.0173)
Post-War	0.0213	0.0175	0.9259
	(0.0023)	(0.0012)	(0.0165)
Garbage	0.0142	0.0276	0.7661
	(0.0041)	(0.0024)	(0.0726)

The table provides the point estimates for the parameters of the trend-stationary consumption process (3.18)*. Bootstrapped standard errors from* 10^6 *simulations are provided in parentheses. For each consumption time series,* $\bar{g} = \mathrm{E}(\Delta c_t)$ *denotes the linear trend. The estimate for the autocorrelation coefficient* ρ_c *is the correlation coefficient of the de-trended levels* $x_t = c_t - g_t$ *and its one-period lag.* σ_ϵ *is the standard deviation of the residuals* $x_t - \rho_c x_{t-1}$*. The long sample consists of all real consumption in the Shiller data set from 1889–2009 and the post-war series from the same data set for 1947–2009. The garbage data of Savov (2011) is available from 1960–2006.*

3.3.3.2 Random Walk Consumption

In addition to the trend-stationary consumption specification, we also consider a model of i.i.d. consumption growth, so consumption has a unit root. Setting $\sigma_\epsilon = 0$ and $\rho_c = 0$, the general specification (3.2) simplifies to the random walk ("RW")

$$\begin{aligned} c_t &= g_t \\ g_t &= \bar{g} + g_{t-1} + \sigma_\nu \nu_t \\ \nu_t &\sim N(0,1) \text{ i.i.d.} \end{aligned} \tag{3.19}$$

For each consumption time series the parameter estimates for the random walk model are $\bar{g} = \mathrm{E}(\Delta c_t)$ and $\sigma_\nu = \sigma(\Delta c_t)$ from Table 3.1, respectively.

3.4 Asset Pricing Implications

We first present the main asset pricing implications of the trend-stationary model and contrast them with those of the random walk model. Subsequently we demonstrate the robustness of the results by examining the implications of various modifications of the consumption process. And finally we analyze the return predictability of the trend-stationary model.

3.4.1 Calibration of the Trend Stationary Model

We calibrate the discount factor β and the risk-aversion coefficient γ so that the asset prices in the trend-stationary model match the empirical values of the risk-free rate and the equity premium. Table 3.7 compares the resulting summary measures of the trend-stationary model to the empirical moments found in the data. The first two rows of the table show the necessary

Table 3.7: Summary Measures of the Trend Stationary Model Compared to the Data

	Long Sample		**Post-War**		**Garbage**	
	Data	Model	Data	Model	Data	Model
γ		7.70		16.5		8.24
β		1.10		1.34		1.08
$E(R_t)$	7.60	7.60	7.92	7.92	7.09	7.09
$\sigma(R_t)$	18.73	23.02	16.60	22.81	15.18	23.92
$E(R_t^f)$	1.97	1.97	1.84	1.84	2.21	2.21
$\sigma(R_t^f)$	5.80	5.85	2.65	5.76	2.60	8.48
EP	5.63	5.63	6.08	6.08	4.87	4.87
SR	0.30	0.24	0.37	0.27	0.32	0.20
$E\,(p_t - d_t)$	3.22	3.54	3.42	3.52	3.51	3.54
$\sigma\,(p_t - d_t)$	0.40	0.42	0.44	0.50	0.40	0.29
AC1 $(p_t - d_t)$	0.92	0.91	0.94	0.93	0.93	0.90

The table compares implied model moments of the trend-stationary model with the empirical moments found in the data. The parameter estimates for the trend-stationary model are given in Table 3.6. The parameters γ and β are calibrated to match the risk-free rate and the equity premium.

values of γ and β for which the asset prices generated by the model match the observed risk-free rate and the equity premium. For the long sample the values are $\gamma = 7.70$ and $\beta = 1.10$, for the post-war series they are $\gamma = 16.5$ and $\beta = 1.34$, and for the garbage data they are $\gamma = 8.24$ and $\beta = 1.08$. Our findings are consistent with a recent strand of literature considering $\beta > 1$. All three pairs of β and γ satisfy the condition (3.12) for the existence of equilibrium when the discount factor exceeds one. The lower volatility and positive autocorrelation in the post-war data leads to a higher implied risk aversion.

For the long sample, the trend-stationary model does a very good job matching the data. The model generates close estimates for the volatility of the risk-free rate as well as all three summary measures for the price-dividend ratio. It slightly overestimates the return volatility for the aggregate consumption claim and thus underestimates the Sharpe ratio. The results are almost as good for the post-war data and the garbage data. In addition to overestimating the return volatility, the model also overestimates the volatility of the risk-free rate, which has been very low in the post-war period.

We next document that all estimates from the trend-stationary model, even those that are a bit off, are much closer to the data than the estimates from the calibrated random walk model. Table 3.8 reports the necessary values of γ and β for which the asset prices generated by the random walk model match the observed risk-free rate and the equity premium.

Table 3.8: Summary Measures of the Random Walk Model Compared to the Data

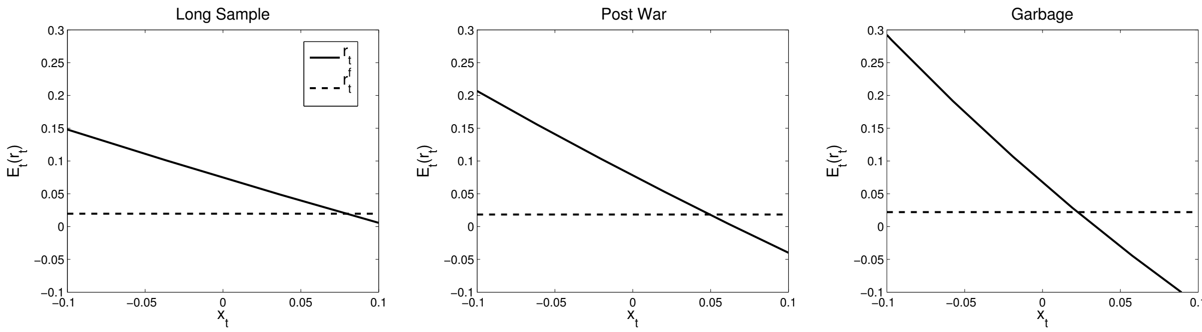
	Long Sample		**Post-War**		**Garbage**	
	Data	RW	Data	RW	Data	RW
γ		43.9		179		57
β		0.7175		0.2475		0.5820
$E(R_t)$	7.60	7.60	7.92	7.92	7.09	7.09
$\sigma(R_t)$	18.73	3.79	16.60	1.94	15.18	3.06
$E(R_t^f)$	1.97	1.97	1.84	1.84	2.21	2.21
$\sigma(R_t^f)$	5.80	0	2.65	0	2.60	0
EP	5.63	5.63	6.08	6.08	4.87	4.87
SR	0.30	1.49	0.37	3.12	0.32	1.59
$E\,(p_t - d_t)$	3.46	2.91	3.42	2.87	3.46	2.89
$\sigma\,(p_t - d_t)$	0.40	0	0.44	0	0.41	0
AC1 $(p_t - d_t)$	0.91	0	0.94	0	0.93	0

The table compares implied model moments of the random walk model with the empirical moments found in the data. The parameter estimates for the trend-stationary model are given in Table 3.1. The parameters γ and β are calibrated to match the risk-free rate and the equity premium.

The necessary values for the risk-aversion coefficient are much larger than for the trend-stationary model; in fact, they are unreasonably large, particularly the value of 179 for the post-war data. The need to match the risk-free rate with such high coefficients of risk-aversion requires very low betas. The random walk model delivers a return volatility (for the aggregate consumption claim) that is much too small. As a result the estimated Sharpe ratios are much too large. The random walk model also underestimates the price-dividend ratio. And as is well-known, the model cannot generate volatility of the risk-free rate and of the price-dividend ratio.

We next provide some intuition for the successful predictive performance of the canonical and parsimonious model with the trend-stationary consumption process. Figure 3.3 shows conditional expected returns of the risky and riskless asset given the state x_t for the three different datasets. We observe that expected excess returns are monotonically decreasing in the state x_t; put differently, expected excess returns are higher than average when the economy is below the trend (negative x_t) and lower than average when it is above the trend (large x_t). This result is in line with the empirical finding that expected returns are large in recessions and low in economic booms. Changes in expected returns are actually the main driver for most of the asset pricing dynamics as Shiller (1981) has shown (instead of changes in expected dividend growth, as had been assumed previously). For a more in-depth analysis of these facts in the context of our model we now decompose the volatility of the price-dividend ratio generated by the model. In fact, we can show that just like in the data most of the variation in the price-dividend ratio is generated by changes in expected returns.

Figure 3.3: Conditional Expected Returns in the Trend-Stationary Model



The graphs show the conditional expected returns of the risky and riskless asset given the state x_t. Results are calculated using the calibrations reported in Table 3.7.

3.4.2 Volatility Tests

It is well known that stock prices move far more than can be explained by changes in expected dividend growth (see Shiller (1981)) and most of the price dynamics are driven by changes in expected returns. To demonstrate that the trend-stationary model captures this fact, we run a volatility test as in Cochrane (1992). The test is based on a first-order approximation of the return identity $R_{t+1} = (P_{t+1} + D_{t+1})/P_t$ which implies that

$$\text{var}(p_t - d_t) \approx \sum_{i=1}^{\infty} \xi^i \text{cov}(p_t - d_t, -r_{t+i}) + \sum_{i=1}^{\infty} \xi^i \text{cov}(p_t - d_t, \Delta d_{t+i}) \tag{3.20}$$

where $\xi = (P/D)/(1 + P/D)$. The linearization is around the point P/D; in our model, this point is the sample mean of the price-dividend ratio. So variations in the price-dividend ratio can only exist, if this ratio has predictive power for either returns or dividend growth or both. Following Campbell and Cochrane (1999), we use 15 years of covariances to approximate the two (infinite) sums on the right-hand side of the expression (3.20). Table 3.9 reports both empirical results and model predictions for all three datasets.

Table 3.9: Variance Decompositions

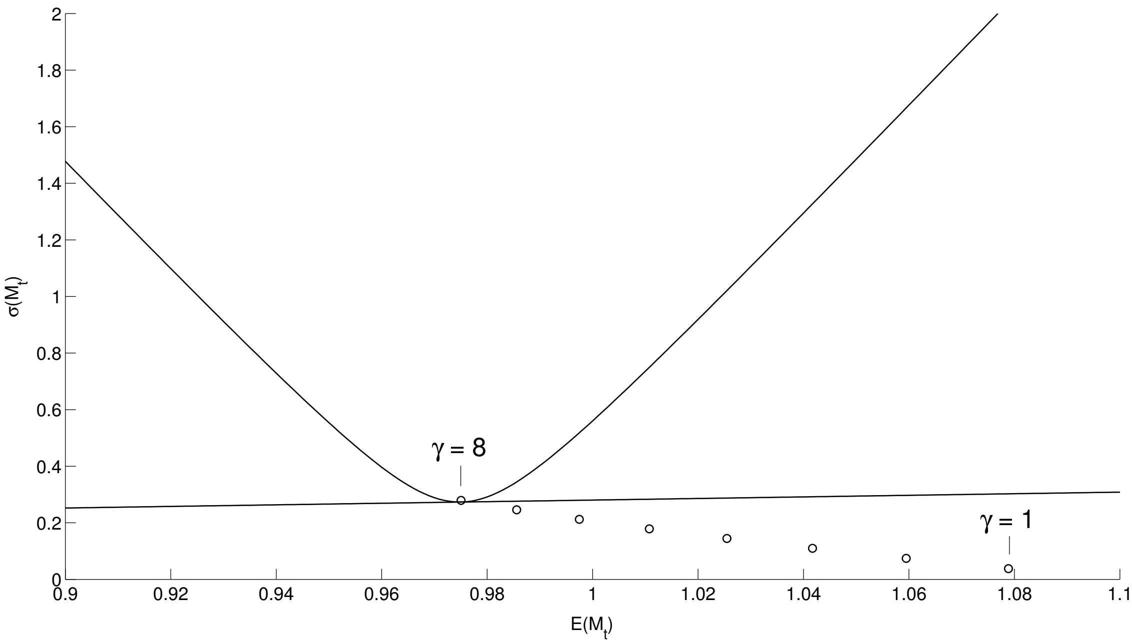
	Data		Model	
	Returns	Dividends	Returns	Dividends
Long Sample	0.7665	-0.0922	0.9719	-0.1283
Post-War	1.1714	-0.2024	0.8597	-0.0532
Garbage	0.6587	0.1316	1.1088	-0.1351

The table shows shares of the variance in the price-dividend ratio that are explained by returns and dividends for the three different datasets. The shares are calculated as $\frac{\sum_{i=1}^{15} \xi^i \text{cov}(p_t - d_t, -r_{t+i})}{\text{var}(p_t - d_t)}$ *and* $\frac{\sum_{i=1}^{15} \xi^i \text{cov}(p_t - d_t, \Delta d_{t+i})}{\text{var}(p_t - d_t)}$*, respectively.*

In the data we find that almost all variation in the price-dividend ratio is driven by changes in expected returns, while the changes attributed to expected dividends are rather small. For the trend-stationary model we find about the same patterns for all three calibrations with most of the variations in the price-dividend ratio coming from changes in expected returns.

A related challenge for asset pricing models is the Hansen and Jagannathan (1991) bound. To generate a high Sharpe ratio, the stochastic discount factor must be very volatile. Figure 3.4 illustrates the relationship between the Hansen-Jagannathan bound and the trend-stationary model SDF for the full sample. For γ around 8 the bound is met. (The corresponding figures for post-war and garbage data are shown in Appendix 3.C.)

Figure 3.4: Hansen-Jagannathan Bounds



The graph shows the mean and standard deviation of the pricing kernel M_t implied by the trend-stationary model as well as the Hansen-Jagannathan bounds for different degrees of risk aversion γ. Each circle represents an increase of γ of one, starting at $\gamma = 1$ in the right lower corner. Parameter estimates for consumption are taken from the long sample, see Table 3.1 with the calibration from Table 3.7, with $\mathrm{E}(\Delta c_t) = 0.020$ *and* $\sigma(\Delta c_t) = 0.0352$, $\rho_c = 0.91$ *and* $\beta = 1.1$.

3.4.3 Robustness Checks on the Consumption Process

We perform a series of robustness checks on the parameters of the consumption process. These checks not only demonstrate the robustness of our results but also help us to develop more intuition for the features of the consumption process that drive the asset pricing implications of the trend-stationary model.

3.4.3.1 Model with both Permanent and Temporary Shocks

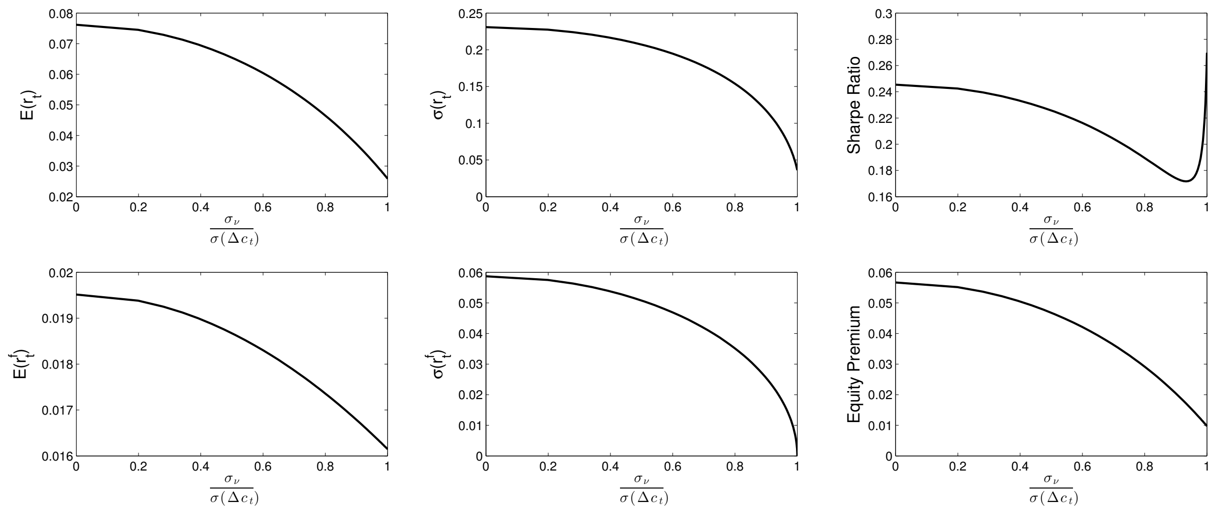
The consumption process (3.18) in the trend-stationary model has a deterministic linear trend with AR(1) deviations; there are no permanent shocks to the growth rate g_t since $\sigma_\nu = 0$. As a first robustness check, we now add a permanent shock to the trend-stationary model. Recall from Equation (3.4) the analytical relationship between the consumption growth volatility $\sigma(\Delta c_t)$ and the standard deviations of the two shocks ϵ_t and ν_t,

$$\sigma(\Delta c_t) = \sqrt{\sigma_\nu^2 + \frac{2}{1+\rho_c}\sigma_\epsilon^2}.$$

We now vary σ_ν and adjust σ_ϵ to hold $\sigma(\Delta c_t)$ constant. We hold the autocorrelation parameter ρ_c constant at its value reported in Table 3.6. In addition, we keep the calibrated values for the risk-aversion coefficient γ and the discount factor β. Note that we deliberately do not recalibrate the model because we want to examine how the summary measures respond to changes in the magnitude of the permanent shock for fixed preferences.

Figure 3.5 shows the influence of the permanent shock σ_ν on several asset pricing characteristics for the long sample. (Appendix 3.C shows the corresponding figures for the post-war and garbage data.) On the horizontal axis we report the share of the permanent shock σ_ν in total consumption growth volatility $\sigma(\Delta c_t)$, that is, the ratio $\sigma_\nu/\sigma(\Delta c_t)$.

Figure 3.5: Asset Pricing Effects of Adding Permanent Shocks σ_ν to the Trend Stationary Model for the Long Dataset



The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the share of the permanent shock σ_ν in the total volatility of consumption growth $\sigma(\Delta c_t)$. (All other volatility comes from the temporary shock σ_ϵ.) Parameter estimates for consumption are taken from the long sample, see Table 3.1 with the calibration from Table 3.7, so $\mathrm{E}(\Delta c_t) = 0.020$ *and* $\sigma(\Delta c_t) = 0.0352$, $\rho_c = 0.91$, $\gamma = 7.7$ *and* $\beta = 1.1$.

At the left end of the horizontal axis, it holds that $\sigma_\nu = 0$, and so all volatility stems from the (temporary) shock ϵ_t. This case corresponds to the pure trend-stationary model, recall the results in Table 3.7. At the right end of the horizontal axis, all consumption volatility comes from the permanent shock, so $\sigma(\Delta c_t) = \sigma_\nu$ and $\sigma_\epsilon = 0$. This case corresponds to the pure random walk model. Observe that all six curves in Figure 3.5 are rather flat as long as the share $\sigma_\nu/\sigma(\Delta c_t)$ is less than 40%. For example, the equity premium remains above 5% in this range. So, the reported summary measures do not change significantly as long as the share of the permanent shock in the consumption volatility remains below 40%. Put differently, the asset pricing implications of the trend-stationary model are very robust to the inclusion of a permanent component in the consumption growth volatility. Temporary shocks to consumption growth, as long as they are sufficiently large, drive the asset pricing implications of the model.

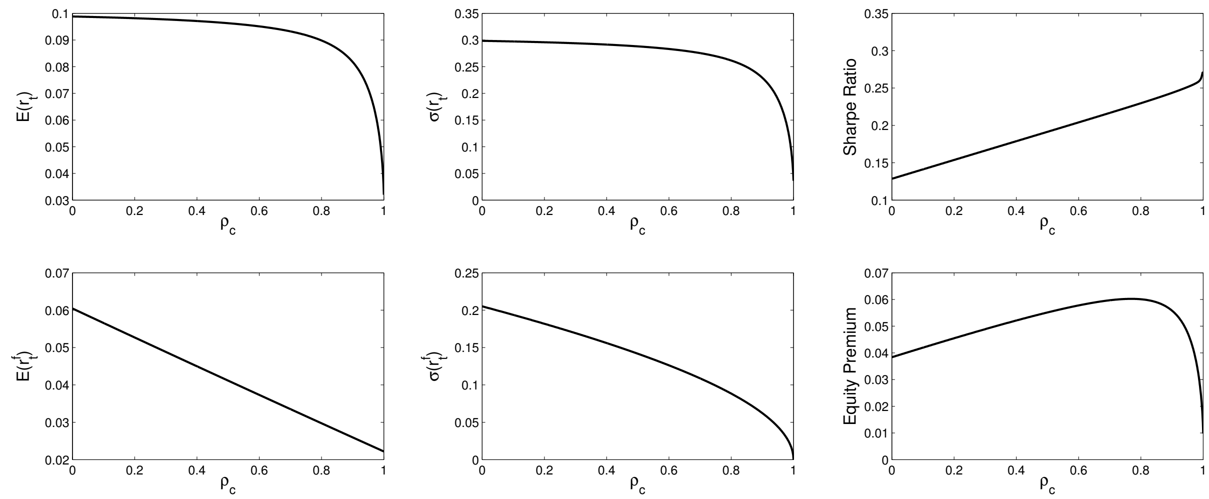
3.4.3.2 Effects of the Autocorrelation Parameter ρ_c

As a second robustness check, we now examine the influence of the autocorrelation parameter ρ_c. Recall that the larger ρ_c, the more persistent is a shock ϵ_t to consumption. (In the extreme case $\rho_c = 1$, the consumption process ceases to be trend-stationary and instead has a unit root and becomes a random walk with drift.) We hold the consumption growth volatility constant; that is, a change in ρ_c implies a change in the standard deviation σ_ϵ to hold $\sigma(\Delta c_t)$ constant according to Equation (3.4). We maintain the calibrated values for the risk-aversion coefficient γ and the discount factor β from Table 3.7.

Figure 3.6 shows the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the expected return of the risk free asset, its volatility and the equity premium as a function of the autocorrelation coefficient in consumption, ρ_c. For all results, the expected value and volatility of consumption growth are fixed at the empirical values of the long sample, so $\mathrm{E}(\Delta c_t) = 0.020$ and $\sigma(\Delta c_t) = 0.0352$. (In Appendix 3.C we show the corresponding figures for the post-war and garbage data series.) We observe that the return volatility of the aggregate consumption claim and the equity premium are very sensitive to changes in the autocorrelation coefficient ρ_c when ρ_c is close to one. The equity premium and both the average and the volatility of the return of the aggregate consumption claim sharply increase for small deviations from the random walk case ($\rho_c = 1$). We also observe that the Sharpe ratio is increasing in the autocorrelation coefficient ρ_c while the equity premium peaks at around 0.8. So, a large level of ρ_c is desired to obtain a high Sharpe ratio while somewhat lower values increase the equity premium and generate more volatility in the asset returns. These findings hold qualitatively for all three datasets, see Figures 3.13 and 3.14 in Appendix 3.C.

In sum, the robustness check with regard to the autocorrelation coefficient ρ_c stresses the importance of trend-stationarity of the consumption process to generate realistic pricing implications in our canonical and parsimonious asset pricing model. Even a modest deviation from the random walk model ($\rho_c \in [0.8, 0.95]$) leads to substantially improved asset pricing implications of the model.

Figure 3.6: Asset Pricing Effects of the Autocorrelation Coefficient ρ_c in Consumption for the Long Dataset



The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the long sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.020$ *and* $\sigma(\Delta c_t) = 0.0352$. *The results are computed using the parameter estimates from Table 3.7 (*$\gamma = 7.7$, $\beta = 1.1$*). (*$\sigma_\nu = 0$*).*

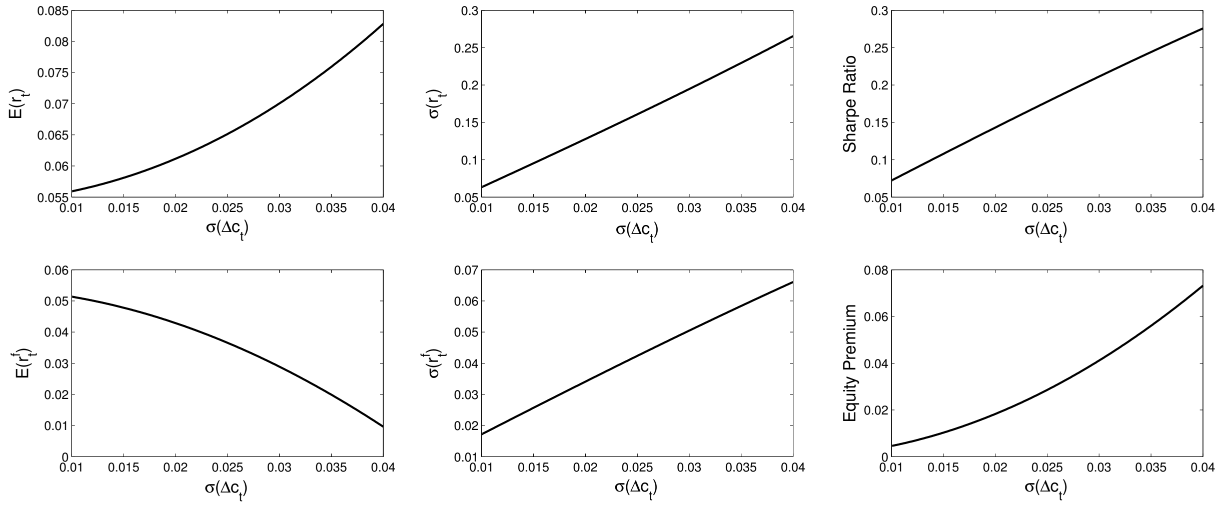
3.4.3.3 Effects of the Consumption Growth Volatility $\sigma(\Delta c_t)$

The volatility of consumption growth, $\sigma(\Delta c_t)$, plays a critical role in the pricing implications of consumption-based asset pricing models. Many papers have pointed out that this volatility in the data is too small to generate non-trivial risk premia in such models, see, among other, Grossman and Shiller (1981) or Mehra and Prescott (1985).

As a third robustness check, we vary $\sigma(\Delta c_t)$ in the interval $[0.01, 0.04]$. This range encompasses the estimated values for all three data sets, see Table 3.1. Once again we use Equation (3.4) to adjust σ_ϵ while holding $\rho_c = 0.91$ and $\sigma_\nu = 0$ constant. We do not recalibrate the model but maintain the values for γ and β from the initial calibration, see Table 3.7. Figure 3.7 shows the usual summary measures as a function of the volatility $\sigma(\Delta c_t) \in [0.01, 0.04]$ in a trend-stationary model based on the long sample, so for the statistical parameters $\bar{g} = \mathrm{E}(\Delta c_t) = 0.02$, $\rho_c = 0.91$, and the model parameters $\gamma = 7.7$ and $\beta = 1.1$. (Appendix 3.C shows the corresponding figures for the post-war and garbage data series.)

While the expected return of the aggregate consumption claim increases considerably with $\sigma(\Delta c_t)$ in the range $[0.01, 0.04]$, the risk-free rate decreases in $\sigma(\Delta c_t)$. As a result, the equity premium increases rather swiftly in the consumption volatility $\sigma(\Delta c_t)$ in the range $[0.01, 0.04]$. Also the return volatilities of the aggregate consumption claim and the risk-free rate are increasing in $\sigma(\Delta c_t)$.

Figure 3.7: Summary Measures for the TS Model as a Function of the Consumption Growth Volatility for the Long Dataset



The graph shows the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the expected risk free rate, its volatility and the equity premium as a function of the standard deviation of consumption growth $\sigma(\Delta c_t)$. The underlying consumption process is from the long sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.020$ and $\rho_c = 0.91$. The results are computed using the parameter estimates from Table 3.7 ($\gamma = 7.7$, $\beta = 1.1$). ($\sigma_\nu = 0$).

Our observations underline the important role of the consumption growth volatility for the asset pricing implications of our canonical and parsimonious model. Rather small increases in the volatility of consumption growth, particularly above a value of 0.03, have a considerable impact on asset prices in the trend-stationary model.

This completes our series of robustness checks on the parameters in the consumption process. In the next step of our analysis, we document the robustness of the asset pricing implications of the trend-stationary model for a different specification of the dividend process. Specifically, we now allow for a time-varying share of dividends in aggregate consumption. Put differently, we explicitly distinguish between the consumption and the dividend process in the economy.

3.4.4 Dividend and Labor Income

We now turn to a generalization of our model to emphasize the robustness of our results with regard to different model specifications. We abandon the assumption of dividends being a fixed fraction of consumption and instead allow for time variation in the shares of financial and labor income in aggregate consumption. This model extension is motivated by recent results in the finance literature. For example, Longstaff and Piazzesi (2004) find that explicitly including dividend payments by the corporate sector in the analysis of macroeconomic cash flows has strong effects on asset prices.[5] Our model specification here is a linear, discrete-time version of the model by Santos and Veronesi (2006) with the assumption of consumption being trend-stationary (instead of consumption growth being stationary). Santos and Veronesi (2006) analyze the impact of time variations in the share of labor income in aggregate consumption on the predictability of stock returns. We build on their results and analyze the effects on expected returns and the equity premium. We follow DeJong and Ripoll (2007) who state that labor endowments, dividends, prices and consumption follow a balanced growth path with an annual common mean growth rate. DeJong and Ripoll (2007) built on the results of Shiller (1981) who assumes that dividends and prices are trend-stationary. DeJong (1992) provides empirical support for this assumption.

In the new model specification, we allow for time variation in the share of financial income in aggregate consumption,

$$C_t = D_t + E_t,$$

where E_t describes all non-dividend income. We write the model in terms of the non-dividend to dividend income ratio to ensure stationarity,

$$\Phi_t = \frac{E_t}{D_t}$$

and, alternatively, in logs

$$\phi_t = e_t - d_t.$$

For the description of equilibria, we employ two state variables, detrended consumption $x_t = c_t - g_t$ and the log non-dividend to dividend income ratio $\phi_t = e_t - d_t$. The claim on the aggregate dividend stream can now be priced in terms of the price-consumption ratio, see Equation (3.7),

$$\begin{aligned} \frac{P_t}{C_t} &= \beta E_t \left\{ e^{(1-\gamma)(g_{t+1}-g_t+x_{t+1}-x_t)} \left(\frac{P_{t+1}}{C_{t+1}} + \frac{D_{t+1}}{C_{t+1}} \right) \right\} \\ &= \beta E_t \left\{ e^{(1-\gamma)(g_{t+1}-g_t+x_{t+1}-x_t)} \left(\frac{P_{t+1}}{C_{t+1}} + \frac{1}{1+e^{\phi_{t+1}}} \right) \right\}, \end{aligned} \tag{3.21}$$

with the resulting one-period return

$$\begin{aligned} R_{t+1} &= \frac{\left(\frac{P_{t+1}}{C_{t+1}} + \frac{D_{t+1}}{C_{t+1}} \right)}{\frac{P_t}{C_t}} \times \frac{C_{t+1}}{C_t} \\ &= \frac{\left(\frac{P_{t+1}}{C_{t+1}} + \frac{1}{1+e^{\phi_{t+1}}} \right)}{\frac{P_t}{C_t}} \times e^{(g_{t+1}-g_t+x_{t+1}-x_t)}. \end{aligned} \tag{3.22}$$

[5] Longstaff and Piazzesi (2004) consider a model where consumption growth and the share of financial income in aggregate consumption are continuous time jump-diffusion processes. They assume a 1% probability of a 10% decline in consumption and a 90% decline in dividends. Their low estimates of the equity premium can be explained by their unit root assumption in consumption. As we see below, the trend-stationary alternative produces significantly larger premia even without the large shocks.

We refer to this specification of the model as the DC (Dividend Claim) model.

3.4.4.1 Estimation of the Dividend Claim Model

To calibrate the model, we use data on observed dividends and consumption. In line with Longstaff and Piazzesi (2004), endowment income is calculated as the difference between aggregate consumption and observed dividends scaled, so that the share of dividends in aggregate consumption is 4% on average.[6] Longstaff and Piazzesi (2004) use a measure they call 'imputed' dividends to account for the fact that corporations tend to artificially smooth dividends over time, so their measure is more volatile than the reported dividends that we use. As the volatility of dividends positively affects the equity premium, our estimate is rather conservative and we restrict ourselves to the lower bounds of expected returns and volatilities. We assume that the state variables x_t and ϕ_t follow two correlated AR(1) processes given by[7]

$$\begin{aligned} x_t &= (1-\rho_c)\mu_x + \rho_c x_{t-1} + \epsilon_{x,t} \\ \phi_t &= (1-\rho_\phi)\mu_\phi + \rho_\phi \phi_{t-1} + \epsilon_{\phi,t} \end{aligned} \tag{3.23}$$

with $\epsilon_{x,t}, \epsilon_{\phi,t} \sim N(0, \boldsymbol{\Sigma_{x,\phi}})$. We estimate the model using simple OLS regressions. Table 3.10 reports the parameter estimates for this model for all three data sets.

Table 3.10: Parameter Estimates for the Pricing of the Dividend Claim

	μ_x	μ_ϕ	ρ_c	ρ_ϕ	σ_c	σ_ϕ	$cov_{c,\phi}$	$\bar{g}$
Long Sample	2.6208	3.5005	0.9100	0.9618	0.0344	0.1134	-0.0002	0.0200
Post-War	2.6336	3.4632	0.9259	0.9841	0.0177	0.0634	-0.0000	0.0213
Garbage	0.2646	3.1873	0.7661	0.9545	0.0269	0.0471	0.0006	0.0142

Parameter estimates using least squares estimation for Equation (3.23) *for the three different datasets.*

3.4.4.2 Pricing the Market Portfolio

We calibrate the discount factor β and the risk-aversion coefficient γ again so that the asset prices in the new model with dividend and labor income match the empirical values of the risk-free rate and the equity premium. Table 3.11 reports the calibrated values of the two parameters as well as the summary measures for the pricing of the dividend claim for the three different data sets.[8]

[6] We also varied this share but it didn't affect our results much.

[7] We also tried a VAR(1,1) specification instead of two correlated AR(1) processes for detrended consumption x_t and the non-dividend to dividend income ratio ϕ_t, but this change in the underlying process has no significant influence on the results. The results for the VAR(1,1) specification are shown in Appendix 3.C.

[8] The corresponding table for the VAR(1,1) specification instead of two correlated AR(1) processes for detrended consumption x_t and the non-dividend to dividend income ratio ϕ_t is shown in Appendix 3.C.

Table 3.11: Summary Measures of the Second Model Compared to the Data

	Long Sample		**Post-War**		**Garbage**	
	Data	Model	Data	Model	Data	Model
γ		7.58		16.2		8.68
β		1.10		1.33		1.08
$E(R_t)$	7.60	7.60	7.92	7.92	7.09	7.09
$\sigma(R_t)$	18.73	23.83	16.60	23.47	15.18	23.59
$E(R_t^f)$	1.97	1.97	1.84	1.84	2.21	2.21
$\sigma(R_t^f)$	5.80	5.78	2.65	5.73	2.60	8.71
EP	5.63	5.63	6.08	6.08	4.87	4.87
SR	0.30	0.24	0.37	0.26	0.32	0.21
$E\left(p_t - d_t\right)$	3.22	3.62	3.42	3.54	3.51	3.52
$\sigma\left(p_t - d_t\right)$	0.40	0.47	0.44	0.51	0.40	0.35
AC1 $\left(p_t - d_t\right)$	0.92	0.92	0.94	0.93	0.93	0.82

The table compares implied model moments of the pricing of the dividend claim with the empirical moments found in the data. The parameter estimates for the trend-stationary dividend and income processes are given in Table 3.10. The model parameters γ and β are calibrated to match the risk-free rate and the equity premium.

Broadly speaking, the results are qualitatively very similar to those for the benchmark model with only the aggregate consumption claim in Section 3.4.1. The calibrated values of the model parameters γ and β are close to those reported in Table 3.7 for the benchmark model. We again find values for the risk aversion coefficient γ that are much smaller than the values commonly found in the economics literature. The new model shows also a similar performance to the benchmark model in matching the empirical moments in the data. The model matches the mean, standard deviation and first order autocorrelation of the log price-dividend ratio for all three data sets pretty well. As in the benchmark model, the model slightly overestimates the volatility of returns to the dividend claim compared to the volatility we find in the data. In sum, the distinction between dividend and labor income in the new model does not lead to better estimates than the benchmark model with only the aggregate consumption claim.

3.4.5 Long-Horizon Predictability

Empirical evidence (Campbell and Shiller (1988b), Fama and French (1988)) suggests that over long horizons the price-dividend ratio predicts cumulative returns: high values predict low future returns, and vice versa. A pure random walk CRRA model of consumption and dividends has a constant price-dividend ratio, so it cannot reproduce such predictions. We test the ability of transient shocks to generate this long-horizon predictability, by comparing the results from simulated model data against the empirical results. In each case, we consider the following specification,

$$R_{t+h} = \hat{\alpha} + \hat{\beta}_1(p_t - d_t) + \epsilon_{t+h},$$

where R_{t+h} is the cumulative return over h periods.

Table 3.12 shows the results for long-horizon regressions. We consider both the trend-stationary consumption claim model of Section 3.4.1, as well as the separate dividend claim model of Section 3.4.4. (The parameters are those specified in Table 3.7 for the pricing of the consumption claim (model CC) and Table 3.11 for the pricing of the dividend claim (model DC).)

Table 3.12: Predictability of Stock Returns

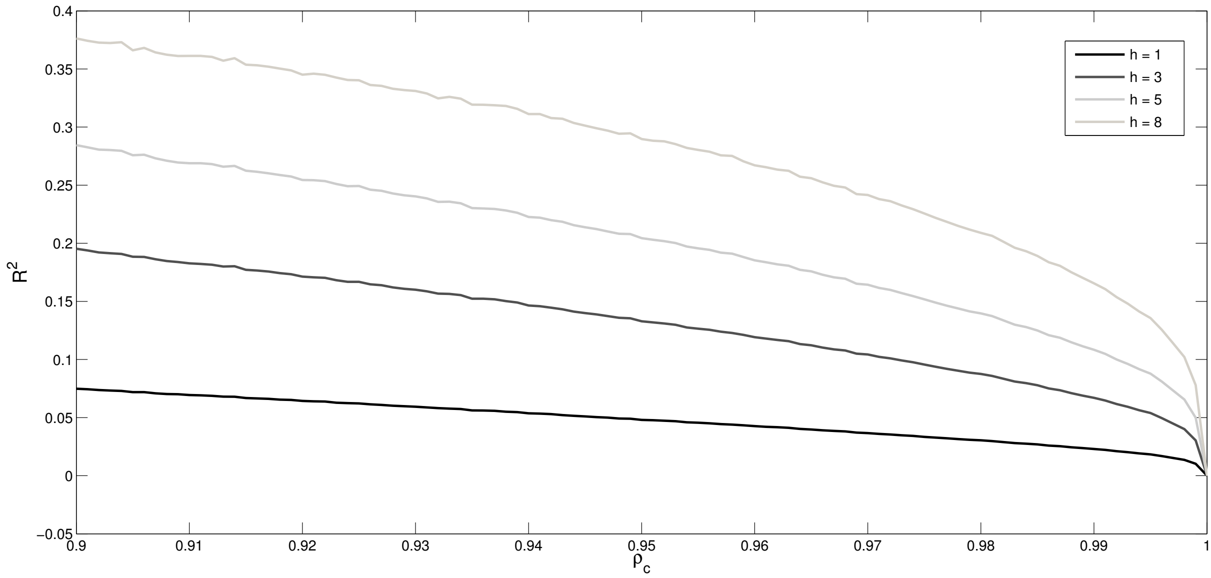
		$h=1$		$h=3$		$h=5$		$h=8$	
		R^2	$\hat{\beta}_1$	R^2	$\hat{\beta}_1$	R^2	$\hat{\beta}_1$	R^2	$\hat{\beta}_1$
	Data	0.0317	-0.0880	0.0644	-0.1962	0.1048	-0.3268	0.1598	-0.5310
L. Samp.	CC	0.0653	-0.1398	0.1725	-0.4433	0.2541	-0.7759	0.3416	-1.3287
	DC	0.0456	-0.1106	0.1192	-0.3479	0.1740	-0.6070	0.2295	-1.0358
	Data	0.1055	-0.1253	0.1920	-0.2800	0.2779	-0.4667	0.3642	-0.7363
P.-War	CC	0.0598	-0.1107	0.1605	-0.3590	0.2394	-0.6454	0.3257	-1.1506
	DC	0.0536	-0.1049	0.1421	-0.3387	0.2100	-0.6074	0.2885	-1.0807
	Data	0.0336	-0.0684	0.0823	-0.1930	0.1385	-0.3493	0.1547	-0.6122
Garb.	CC	0.1351	-0.3099	0.3076	-0.8237	0.4023	-1.2278	0.4706	-1.7058
	DC	0.1359	-0.2594	0.3104	-0.6925	0.4022	-1.0356	0.4649	-1.4397

The table shows the R^2 statistic and the slope coefficient $\hat{\beta}_1$ of regressing cumulative stock returns on the log price-dividend ratio for different time horizons h (in years) for the three different datasets. Parameter estimates for the models are chosen as in Table 3.7 and Table 3.11.

In the data we find the standard patterns documented by Campbell and Shiller (1988c). Coefficients are negative, so high price-dividend ratios today imply low future returns. The R^2 are low for short-term predictions but grow rapidly for longer horizons. The trend-stationary models qualitatively reproduces these findings. For all three data series, the model produces R^2 statistics that are increasing in the horizon of the prediction. The R^2 statistics are somewhat too large for the long sample and the garbage data series and slightly too low for the post-war data.

To illustrate how temporary shocks help generate return predictability, we illustrate in Figure 3.8 the influence of the autocorrelation coefficient in consumption, ρ_c, on the predictability of stock returns (as measured by R^2). Results are shown for the long dataset. When changing ρ_c we adjust σ_ϵ to fix $\sigma(\Delta c_t)$ at its empirical estimate. At $\rho_c = 1$, the random walk case, return predictability disappears, while it grows quickly as ρ_c decreases below 1. Corresponding graphs for the post-war and the garbage dataset are shown in Appendix 3.C.

Figure 3.8: Influence of the Autocorrelation Coefficient ρ_c on the Predictability of Stock Returns for the Long Dataset



The graph shows the predictability of stock returns R^2 for different forecasting horizons h as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the long sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.020$ *and* $\sigma(\Delta c_t) = 0.0352$. *The results are computed using the parameter estimates from Table 3.7 ($\gamma = 7.7$, $\beta = 1.1$). ($\sigma_\nu = 0$).*

3.5 Conclusion

The macroeconomic data on the U.S. economy is consistent with temporary shocks that dissipate. In this paper we have documented the significance of temporary shocks to consumption for consumption-based asset pricing. For a simple theoretical model where some consumption shocks are permanent and other shocks are not, the temporary shock dominates the dynamics of asset prices. We also numerically calibrate a standard CRRA consumption model with temporary shocks using U.S. data, and find that we can match the equity premium and risk-free rate with moderate levels of risk aversion. This same model also generates many of the features of stock prices that have been considered puzzles. Asset prices in the model are very volatile, consistent with the excess volatility puzzle (they are actually somewhat more volatile than in the actual data). Consistent with the variance decomposition of Campbell and Shiller, changes in the model's price-dividend ratio are largely driven by changes in model expected returns, which are large, rather than changes in expected dividends, which are small. High price-dividend ratios today predict low expected returns in the future, consistent with the return predictability literature.

From time-series evidence alone, it is difficult to distinguish between shocks to consumption that are permanent from those that are very persistent, or even a mix of temporary and permanent shocks. Our findings show that the presence of temporary shocks has a dramatic impact on asset pricing dynamics, making many of the empirical aggregate stock market puzzles less puzzling. Permanent shocks have a much smaller impact. Even in a model with both temporary and permanent shocks, the temporary shocks generate most of the dynamics. Together, these results suggest that temporary shocks to consumption are an important, yet largely overlooked, mechanism to explain puzzles in asset pricing.

3.A Analytical Results

We provide proofs for the theoretical results stated in the paper.

3.A.1 The Unconditional Moments of Consumption Growth

We derive the unconditional moments of consumption growth (3.3)–(3.5). Log consumption growth is given by

$$\Delta c_t = \bar{g} + (\rho_c - 1)(x_{t-1}) + \sigma_\epsilon \epsilon_t + \sigma_\nu \nu_t, \tag{3.24}$$

with $\mathrm{E}(x_t) = 0$ and $\mathrm{Var}(x_t) = \sigma_\epsilon^2/(1-\rho_c^2)$. Thus,

$$\mathrm{E}(\Delta c_t) = \bar{g} \quad \text{and} \quad \sigma(\Delta c_t) = \sqrt{\sigma_\nu^2 + \frac{2}{1+\rho_c}\sigma_\epsilon^2}.$$

The auto-covariance of log consumption growth $\mathrm{cov}(\Delta c_t, \Delta c_{t-1})$ is given by

$$\begin{aligned} \mathrm{cov}(\Delta c_t, \Delta c_{t-1}) &= E(\Delta c_t \Delta c_{t-1}) - E(\Delta c_t)^2 \\ &= \bar{g}^2 + (\rho_c - 1)^2 E(x_t x_{t-1}) + (\rho_c - 1)E(x_t \sigma_\epsilon \epsilon_t) - \bar{g}^2 \\ &= \frac{\rho_c - 1}{\rho_c + 1}\sigma_\epsilon. \end{aligned} \tag{3.25}$$

The resulting first-order autocorrelation is $\mathrm{AC1}(\Delta c_t) = \frac{\mathrm{cov}(\Delta c_t, \Delta c_{t-1})}{\sigma(\Delta c_t)^2}$.

3.A.2 Equilibria in Growth Economies with Discount Factors $\beta > 1$

We derive Condition (3.12) under which the agent's expected utility,

$$\frac{1}{1-\gamma} E_t \left[\sum_{k=0}^{\infty} \beta^k C_{t+k}^{1-\gamma} \right],$$

remains finite for any starting values g_t, x_t of the consumption process (3.2). The two processes g_t and x_t are independent and thus

$$E_t\left(C_{t+k}^{1-\gamma}\right) = E_t\left(e^{(1-\gamma)(g_{t+k}+x_{t+k})}\right) = E_t\left(e^{(1-\gamma)g_{t+k}}\right) E_t\left(e^{(1-\gamma)x_{t+k}}\right).$$

Trivially, $E_t\left(C_{t+k}^{1-\gamma}\right)$ is positive for all $k = 0, 1, 2, \ldots$. However, we can find a much tighter lower bound. Note that $E_t(g_{t+k}) = g_t + \bar{g}k$ and so due to the convexity of the exponential function and Jensen's inequality,

$$E_t\left(e^{(1-\gamma)g_{t+k}}\right) \geq e^{(1-\gamma)(g_t + \bar{g}k)}.$$

Similarly, note that $E_t(x_{t+k}) = \rho_c^k x_t$ and so again by Jensen's inequality,

$$E_t\left(e^{(1-\gamma)x_{t+k}}\right) \geq e^{(1-\gamma)\rho_c^k x_t}.$$

And so we obtain the lower bound

$$E_t\left(C_{t+k}^{1-\gamma}\right) \geq e^{(1-\gamma)(g_t+\rho_c^k x_t)} e^{(1-\gamma)(\bar{g}k)}.$$

Next, for $1-\gamma < 0$ we obtain the inequality

$$\frac{1}{1-\gamma}\sum_{k=0}^{\infty} E_t\left[\beta^k C_{t+k}^{1-\gamma}\right] \leq \frac{1}{1-\gamma}\sum_{k=0}^{\infty}\left(e^{(1-\gamma)(g_t+\rho_c^k x_t)}\right)\left(\beta e^{(1-\gamma)\bar{g}}\right)^k.$$

The first exponential term in the product on the right-hand side has a uniform upper bound for all $k = 0, 1, 2, \ldots$, since $\rho_c \in [0,1]$. Therefore, the right-hand side is absolutely convergent as long as

$$\beta e^{(1-\gamma)\bar{g}} < 1.$$

Under this condition, the series on the left-hand side of the inequality is absolutely convergent as well. Therefore, Fubini's Theorem implies that

$$\frac{1}{1-\gamma} E_t\left[\sum_{k=0}^{\infty}\beta^k C_{t+k}^{1-\gamma}\right] = \frac{1}{1-\gamma}\sum_{k=0}^{\infty} E_t\left[\beta^k C_{t+k}^{1-\gamma}\right].$$

This completes the proof that an equilibrium exists if $\gamma > 1$ and $\beta e^{(1-\gamma)\bar{g}} < 1$.

3.A.3 Proof of Theorem 3

We prove Theorem 3 by deriving the expressions (3.14)–(3.18). Recall the consumption process (3.13),

$$\begin{aligned} c_t &= g_t + \sigma_\epsilon \epsilon_t \\ g_t &= \bar{g} + g_{t-1} + \sigma_\nu \nu_t \\ \epsilon_t, \nu_t &\sim N(0,1) \text{ i.i.d.} \end{aligned} \tag{3.26}$$

with the mean growth rate $\bar{g}$ of consumption. The time t conditional expectation of $C_{t+1}^a = e^{ac_t}$ is then given by

$$E_t(C_{t+1}^a) = e^{a(\bar{g}+g_t)+\frac{1}{2}a^2\sigma_\epsilon^2+\frac{1}{2}a^2\sigma_\nu^2}.$$

3.A.3.1 Return and Volatility of the Risk Free Rate

Equation (3.6) determines the price of risk-free one-period bond and implies

$$\begin{aligned} P_t^f &= \beta E_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right\} \\ &= \beta e^{-\gamma \bar{g} + \gamma \sigma_\epsilon \epsilon_t + \frac{1}{2}\gamma^2 \sigma_\epsilon^2 + \frac{1}{2}\gamma^2 \sigma_\nu^2}. \end{aligned}$$

The resulting expected one-period return is then

$$E(R_t^f) = \frac{1}{\beta} e^{\gamma \bar{g} - \frac{1}{2}\gamma^2 \sigma_\nu^2}.$$

For the volatility of the risk-free rate we first derive

$$E((R_t^f)^2) = \frac{1}{\beta^2} e^{2\gamma \bar{g} - \gamma^2 \sigma_\nu^2 + \gamma^2 \sigma_\epsilon^2}$$

to obtain

$$\sigma(R_t^f) = \frac{1}{\beta} e^{\gamma \bar{g} - \frac{1}{2}\gamma^2 \sigma_\nu^2} \left(e^{\gamma^2 \sigma_\epsilon^2} - 1 \right)^{\frac{1}{2}}.$$

3.A.3.2 Return of the Infinitely-Lived Asset

Rewriting Equation (3.7) we obtain

$$\frac{P_t}{C_t^\gamma} = \beta E_t \left\{ \frac{P_{t+1}}{C_{t+1}^\gamma} + C_{t+1}^{1-\gamma} \right\},$$

which, after dividing both sides by $e^{(1-\gamma)g_t}$, yields

$$\frac{P_t}{C_t^\gamma e^{(1-\gamma)g_t}} = \beta E_t \left\{ \frac{P_{t+1}}{C_{t+1}^\gamma e^{(1-\gamma)g_{t+1}}} e^{(1-\gamma)(\bar{g} + \sigma_\nu \nu_{t+1})} + \frac{C_{t+1}^{1-\gamma}}{e^{(1-\gamma)g_t}} \right\}.$$

Now we employ a guess-and-verify approach. Define $X \doteq \frac{P_t}{C_t^\gamma e^{(1-\gamma)g_t}}$ and assume that X is constant. Then

$$X = \beta E_t \left\{ X e^{(1-\gamma)(\bar{g} + \sigma_\nu \nu_{t+1})} + \frac{C_{t+1}^{1-\gamma}}{e^{(1-\gamma)g_t}} \right\}$$

which implies

$$X = \frac{\beta e^{(1-\gamma)\bar{g} + \frac{1}{2}(1-\gamma)^2 (\sigma_\epsilon^2 + \sigma_\nu^2)}}{1 - \beta e^{(1-\gamma)\bar{g} + \frac{1}{2}(1-\gamma)^2 \sigma_\nu^2}}.$$

Note that the right-hand side is indeed independent of time and thus a constant. Using the definition of X, we can express the one-period return of the infinitely-lived asset as

$$\begin{aligned} R_{t+1} &= \frac{P_{t+1}+D_{t+1}}{P_t} - 1 \\ &= \frac{C_{t+1}^{\gamma}}{C_t^{\gamma}} e^{(1-\gamma)(\bar{g}+\sigma_\nu \nu_{t+1})} + \frac{C_{t+1}}{X C_t^{\gamma}} e^{-(1-\gamma)g_t}. \end{aligned}$$

Observing that

$$\frac{C_{t+1}^{\gamma}}{C_t^{\gamma}} = e^{\gamma(\bar{g}+\sigma_\nu \nu_{t+1}+\sigma_\epsilon \epsilon_{t+1}-\sigma_\epsilon \epsilon_t)}$$

and taking conditional expectations yields

$$E_t(R_{t+1}) = e^{\bar{g}+\frac{1}{2}\sigma_\nu^2+\frac{1}{2}\gamma^2\sigma_\epsilon^2-\gamma\sigma_\epsilon\epsilon_t} + \frac{e^{\bar{g}+\frac{1}{2}\sigma_\nu^2+\frac{1}{2}\sigma_\epsilon^2-\gamma\sigma_\epsilon\epsilon_t}}{X}$$

which in turn results in an unconditional expected return of

$$E(R_{t+1}) = e^{\bar{g}+\frac{1}{2}\sigma_\nu^2+\gamma^2\sigma_\epsilon^2} + \frac{1-\beta e^{(1-\gamma)\bar{g}+\frac{1}{2}(1-\gamma)^2\sigma_\nu^2}}{\beta} e^{\gamma\bar{g}+\gamma\sigma_\nu^2-\frac{1}{2}\gamma^2\sigma_\nu^2+\gamma\sigma_\epsilon^2}.$$

Finally, Equation (3.18) follows from the second moment of the return,

$$E(R_{t+1}^2) = e^{4\gamma^2\sigma_\epsilon^2+2\bar{g}+2\sigma_\nu^2} + \frac{1}{X^2} e^{2\gamma^2\sigma_\epsilon^2+2\sigma_\epsilon^2+2\bar{g}+2\sigma_\nu^2} + \frac{2}{X} e^{2.5\gamma^2\sigma_\epsilon^2+\gamma\sigma_\epsilon^2+0.5\sigma_\epsilon^2+2\bar{g}+2\sigma_\nu^2},$$

and the standard deviation $\sigma(R_t) = \left(E(R_t^2) - E(R_t)^2\right)^{\frac{1}{2}}$ of the aggregate consumption claim. This completes the derivations of the statements in Theorem 3.

3.B Numerical Solution Method

We briefly describe the numerical solution approach for the baseline version of the economic model. The solution method relies on quadrature and projection techniques, see Judd (1992) and Judd (1998). A detailed description of the solution method is given in Chapter 4 of this thesis.

Equations (3.8) and (3.10) determine the risk-free rate in the baseline model. Similarly, Equations (3.9) and (3.11) determine the return of the long-lived risky asset. Both the price P_t^f of the riskless one-period asset and the price-consumption ratio P_t/C_t of the infinitely-lived risky asset depend on the detrended consumption x_t which serves as the endogenous state variable for the model. The state space is the interval given by ± 6 standard deviations around the unconditional mean of detrended consumption. On this state space we approximate the functions P_t^f and P_t/C_t by Chebyshev polynomials of degree 18 and use the Galerkin projection method to find the best approximation. (We checked approximations up to degree 32 on a state space as wide as ± 20 standard deviations around the steady state and obtained effectively identical solutions.) We solve the integrals arising due to the Galerkin projection by Gauss-Chebyshev quadrature.[9]

The conditional expectations operator E_t in Equations (3.10) and (3.11) is a two-dimensional integral over x_{t+1} and g_{t+1}, that is, an integral over ϵ_{t+1} and ν_{t+1}. We approximate this double integral by two-dimensional Gauss-Hermite quadrature. As a result we obtain two linear systems of equations in the coefficients of the Chebyshev polynomials which are straightforward to solve. The respective solutions yield approximations of the functions P_t^f and P_t/C_t on the state space.

Finally we can compute returns. For the average risk-free rate and its standard deviation, we integrate $1/P_t^f$ using Gauss-Hermite quadrature on the unconditional distribution of the state variable. Equation (3.9) provides the conditional return of the aggregate consumption claim given the state x_t. We first compute conditional expected returns by a two-dimensional Gauss-Hermite quadrature, again over x_{t+1} and g_{t+1} on the Gauss-Hermite nodes of the unconditional distribution of x_t. Once the conditional returns have been computed, we can calculate unconditional returns by one-dimensional Gauss-Hermite quadrature. (Alternatively, we could simulate the economy for hundreds of thousands of periods and report simulated returns. That approach yields the same results but requires much longer running times.)

For the return predictability, we simulate the underlying processes over 1,000,000 periods and calculate the corresponding prices and returns.

3.C Additional Tables and Graphs

In Section 3.4.4.1 we assume that the state variables x_t and ϕ_t follow two correlated AR(1) processes. Alternatively, we also tried a VAR(1,1) specification for the two state variables. For such a specification, Table 3.13 reports the summary measures of asset prices for the dividend claim model.

Figures 3.4–3.8 show the asset pricing effects of parameter changes for the long sample. For completion, we include the corresponding figures for the post-war data and the garbage series.

Figure 3.4 in Section 3.4.2 shows the Hansen-Jagannathan bounds for the long sample. Figures 3.9 and 3.10 are the corresponding figures for the post-war and the garbage data, respectively.

[9] We also solved the model by using the discretization methods of Tauchen (1986) or Tauchen and Hussey (1991). We obtain almost identical results but find that these methods require many more nodes than the Galerkin method to deliver a good approximation of the solution.

Table 3.13: Summary Measures of the Second Model Compared to the Data – VAR(1,1) Specification

	Long Sample		**Post-War**		**Garbage**	
	Data	Model	Data	Model	Data	Model
γ		7.00		17.90		8.70
β		1.10		1.37		1.08
$E(R_t)$	7.60	7.60	7.92	7.92	7.09	7.09
$\sigma(R_t)$	18.73	26.72	16.60	22.31	15.18	23.57
$E(R_t^f)$	1.97	1.97	1.84	1.84	2.21	2.21
$\sigma(R_t^f)$	5.80	6.82	2.65	7.41	2.60	8.88
EP	5.63	5.63	6.08	6.08	4.87	4.87
SR	0.30	0.21	0.37	0.27	0.32	0.21
$E\,(p_t - d_t)$	3.22	3.87	3.42	3.47	3.51	3.53
$\sigma\,(p_t - d_t)$	0.40	0.55	0.44	0.50	0.40	0.36
AC1 $(p_t - d_t)$	0.92	0.93	0.94	0.94	0.93	0.83

The table compares implied model moments of the pricing of the dividend claim with the empirical moments found in the data. The table shows the same summary statistics as table 3.11, but instead of two correlated AR(1) processes we assume an VAR(1,1) process for detrended consumption x_t and the non dividend to dividend income ratio ϕ_t. Again, the model parameters γ and β are calibrated to match the risk-free rate and the equity premium.

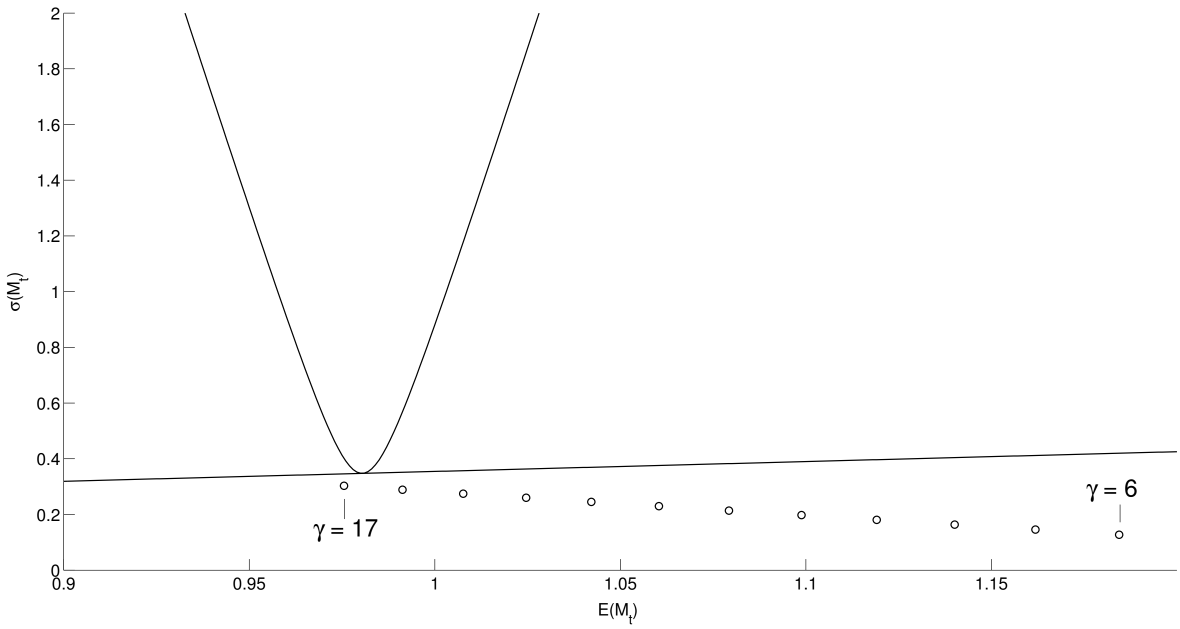
Figure 3.5 in Section 3.4.3.1 shows the effects of adding a permanent shock to the trend-stationary model for the long sample. Figures 3.11 and 3.12 are the corresponding figures for the post-war and the garbage data, respectively.

Figure 3.6 in Section 3.4.3.2 shows the effects of changes in the autocorrelation coefficient ρ_c for the long sample. Figures 3.13 and 3.14 are the corresponding figures for the post-war and the garbage data, respectively.

Figure 3.7 in Section 3.4.3.3 shows the effects of changes in the consumption growth volatility $\sigma(\Delta c_t)$ for the long sample. Figures 3.15 and 3.16 are the corresponding figures for the post-war and the garbage data, respectively.

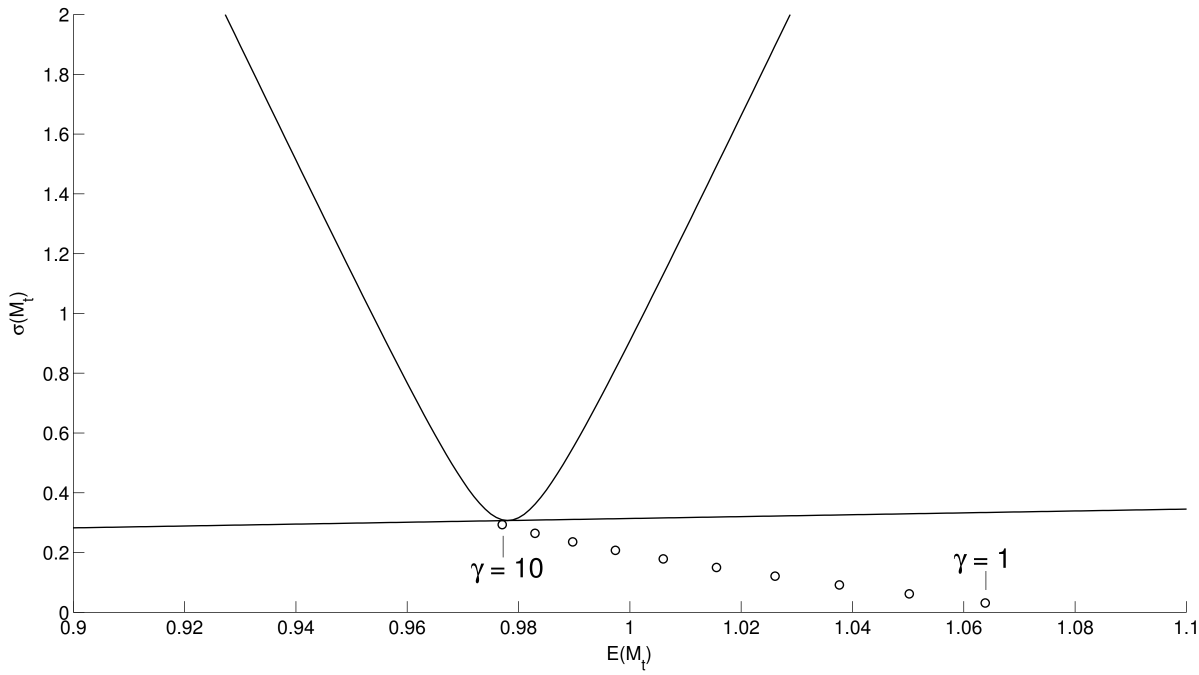
Figure 3.8 in Section 3.4.5 shows the influence of the autocorrelation parameter ρ_c on the predictability of returns for the long sample. Figures 3.17 and 3.18 are the corresponding figures for the post-war and the garbage data, respectively.

Figure 3.9: Hansen-Jagannathan Bounds for Post-War Data



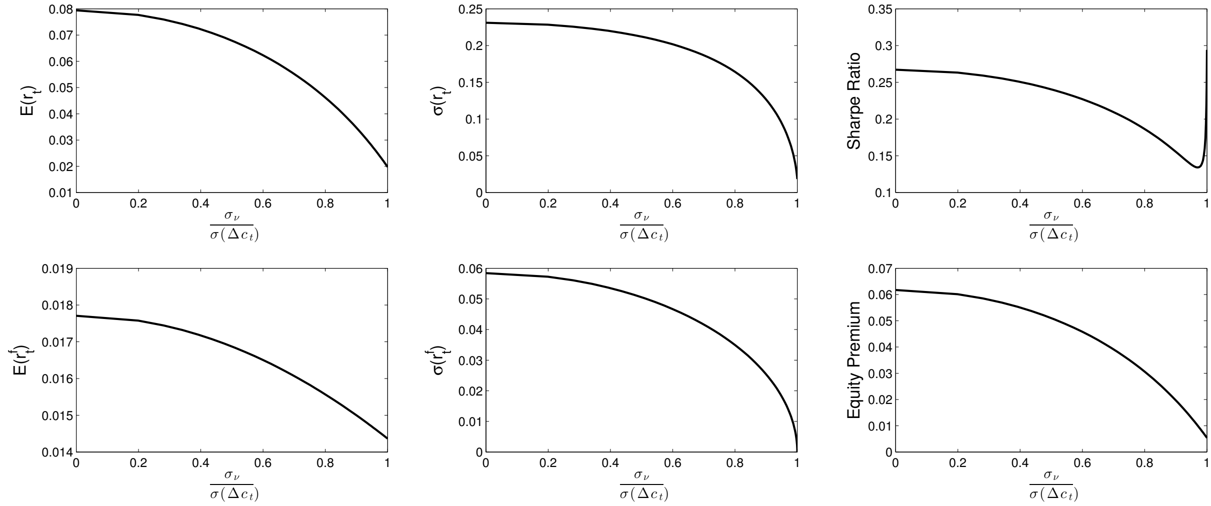
The graph shows the mean and standard deviation of the pricing kernel M_t implied by the trend-stationary model as well as the Hansen-Jagannathan bounds for different degrees of risk aversion γ. Each circle represents an increase of γ of one, starting in the right lower corner. Parameter estimates for consumption are taken from the post-war sample, see Table 3.1 with the calibration from Table 3.7, so $\mathrm{E}(\Delta c_t) = 0.0213$ *and* $\sigma(\Delta c_t) = 0.0180$, $\rho_c = 0.9259$ *and* $\beta = 1.34$.

Figure 3.10: Hansen-Jagannathan Bounds for Garbage Data



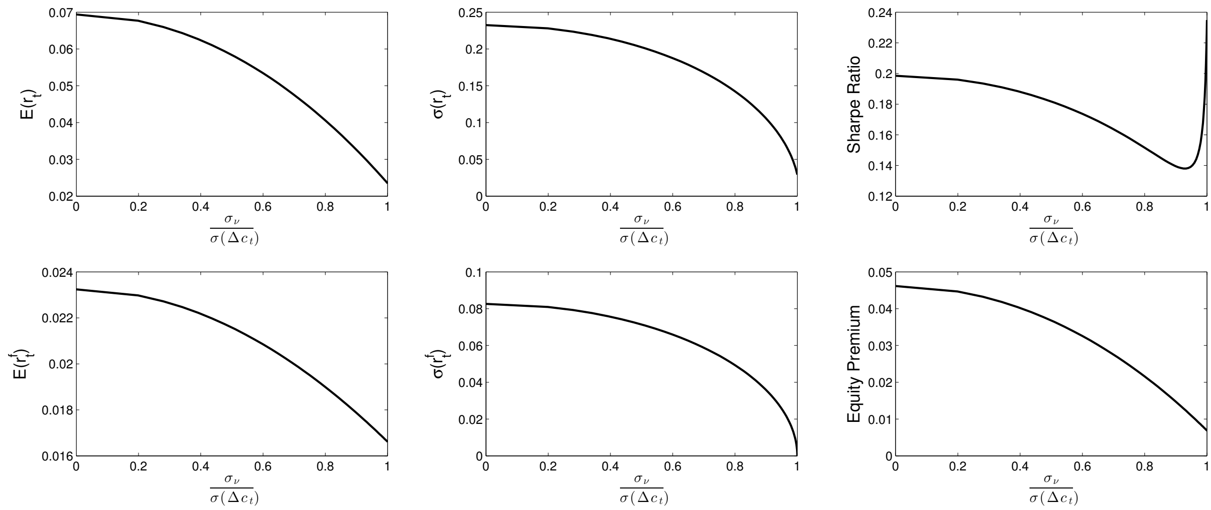
The graph shows the mean and standard deviation of the pricing kernel M_t implied by the trend-stationary model as well as the Hansen-Jagannathan bounds for different degrees of risk aversion γ. Each circle represents an increase of γ of one, starting in the right lower corner. Parameter estimates for consumption are taken from the garbage sample, see Table 3.1 with the calibration from Table 3.7, so $\mathrm{E}(\Delta c_t) = 0.0142$ *and* $\sigma(\Delta c_t) = 0.0286$, $\rho_c = 0.7661$ *and* $\beta = 1.08$.

Figure 3.11: Asset Pricing Effects of Adding Permanent Shocks σ_ν to the Trend Stationary Model for the Post-War Dataset



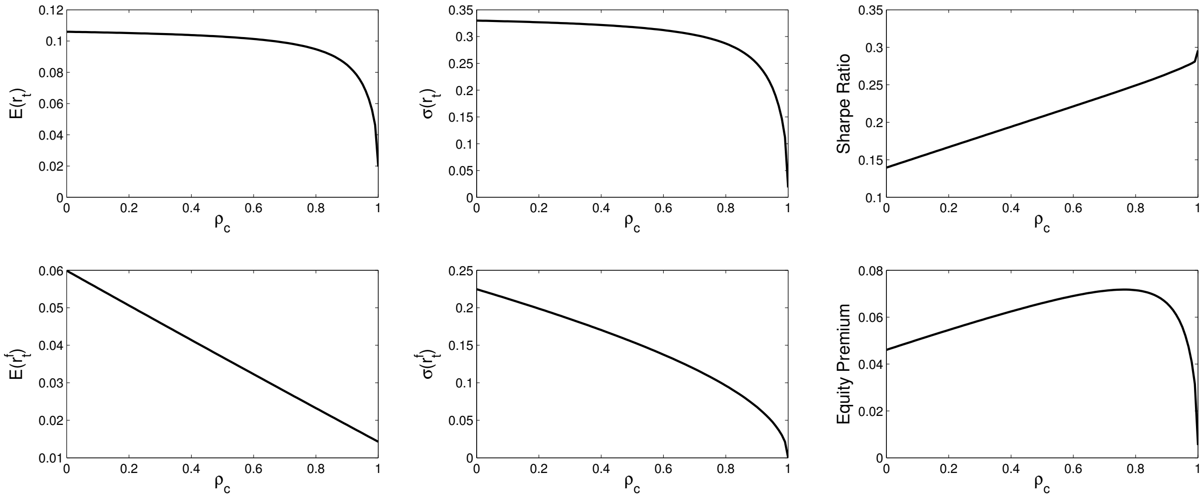
The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the share of the permanent shock σ_ν in the total volatility of consumption growth $\sigma(\Delta c_t)$ (All other volatility comes from the temporary shock σ_ϵ). Parameter estimates for consumption are taken from the post-war sample, see Table 3.1 with the calibration from Table 3.7, so $\mathrm{E}(\Delta c_t) = 0.0213$ *and* $\sigma(\Delta c_t) = 0.018$, $\rho_c = 0.9259$, $\gamma = 16.5$ *and* $\beta = 1.34$.

Figure 3.12: Asset Pricing Effects of Adding Permanent Shocks σ_ν to the Trend Stationary Model for the Garbage Dataset



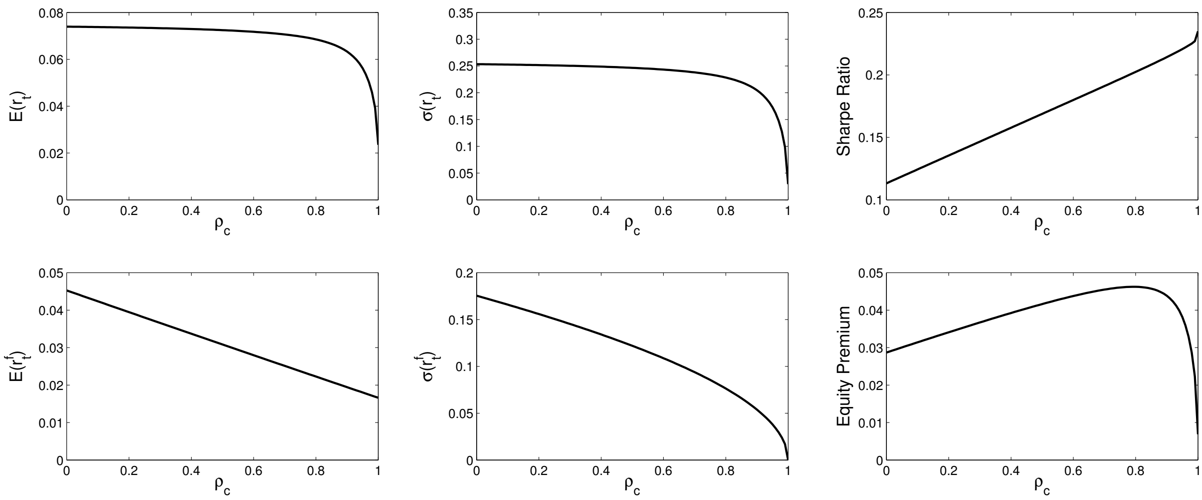
The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the share of the permanent shock σ_ν in the total volatility of consumption growth $\sigma(\Delta c_t)$ (All other volatility comes from the temporary shock σ_ϵ). Parameter estimates for consumption are taken from the garbage sample, see Table 3.1 with the calibration from Table 3.7, so $\mathrm{E}(\Delta c_t) = 0.0142$ *and* $\sigma(\Delta c_t) = 0.0286$, $\rho_c = 0.7661$, $\gamma = 8.24$ *and* $\beta = 1.08$.

Figure 3.13: Asset Pricing Effects of the Autocorrelation Coefficient ρ_c in Consumption for the Post-War Dataset



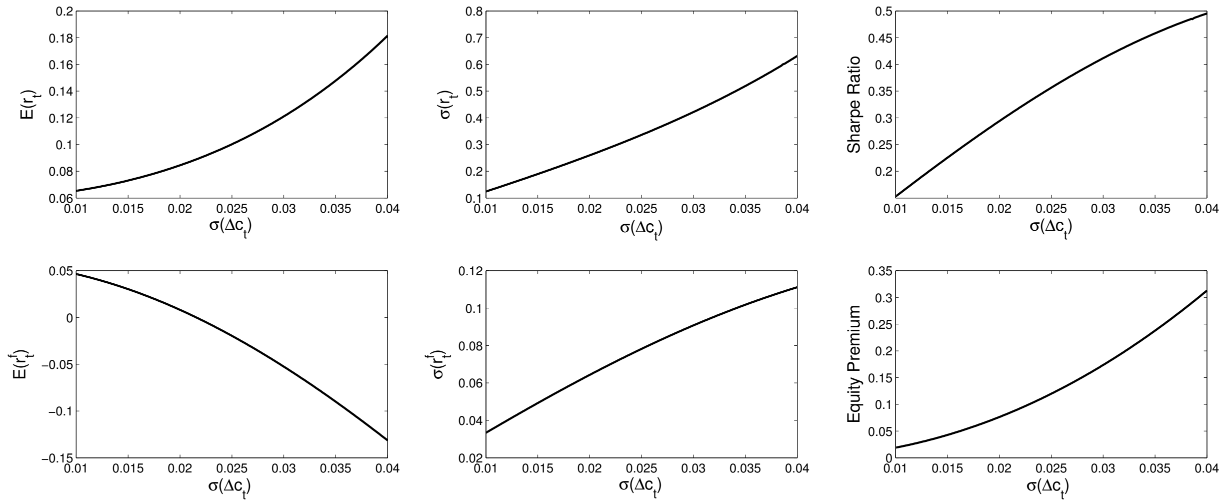
The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the post-war sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.0213$ *and* $\sigma(\Delta c_t) = 0.0180$. *The results are computed using the parameter estimates from Table 3.7 ($\gamma = 16.5$, $\beta = 1.34$). ($\sigma_\nu = 0$).*

Figure 3.14: Asset Pricing Effects of the Autocorrelation Coefficient ρ_c in Consumption for the Garbage Dataset



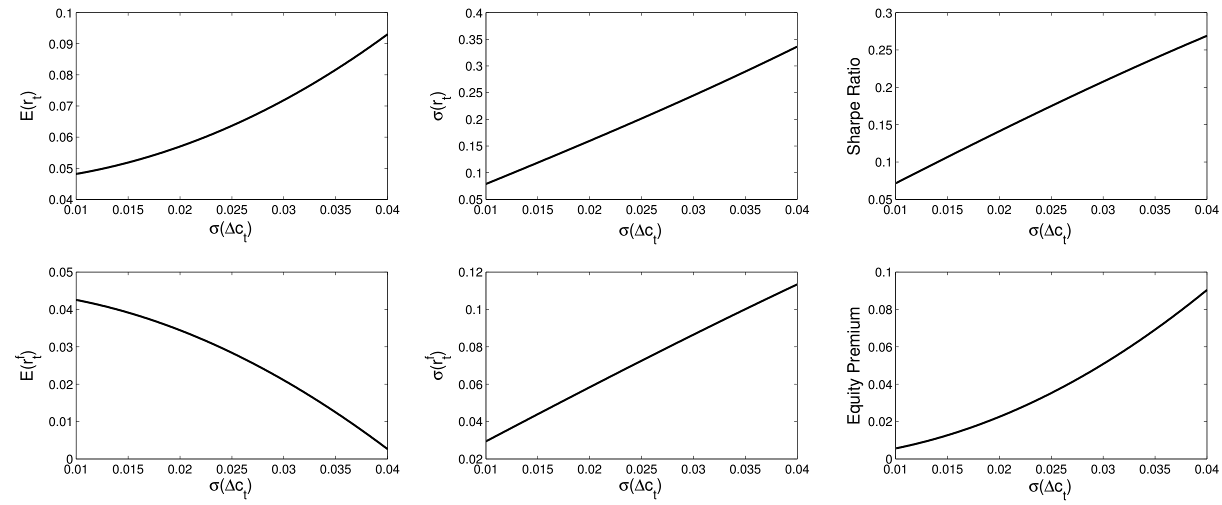
The graphs show the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the risk-free rate, its volatility, and the equity premium as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the garbage data sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.0142$ *and* $\sigma(\Delta c_t) = 0.0286$. *The results are computed using the parameter estimates from Table 3.7 ($\gamma = 8.24$, $\beta = 1.08$). ($\sigma_\nu = 0$).*

Figure 3.15: Summary Measures for the TS Model as a Function of the Consumption Growth Volatility for the Post-War Sample



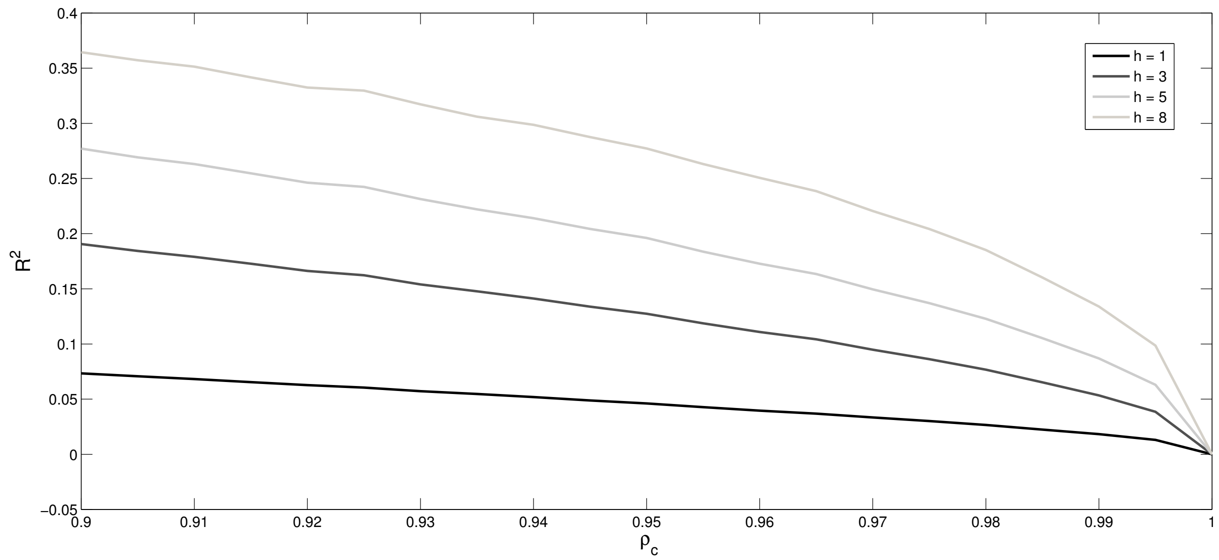
The graph shows the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the expected risk free rate, its volatility and the equity premium as a function of the standard deviation of consumption growth $\sigma(\Delta c_t)$. *The underlying consumption process is from the post-war sample, see Table 3.1, so* $\mathrm{E}(\Delta c_t) = 0.0213$ *and* $\rho_c = 0.9259$. *The results are computed using the parameter estimates from Table 3.7 (*$\gamma = 16.5$, $\beta = 1.34$*). (*$\sigma_\nu = 0$*).*

Figure 3.16: Summary Measures for the TS Model as a Function of the Consumption Growth Volatility for the Garbage Sample



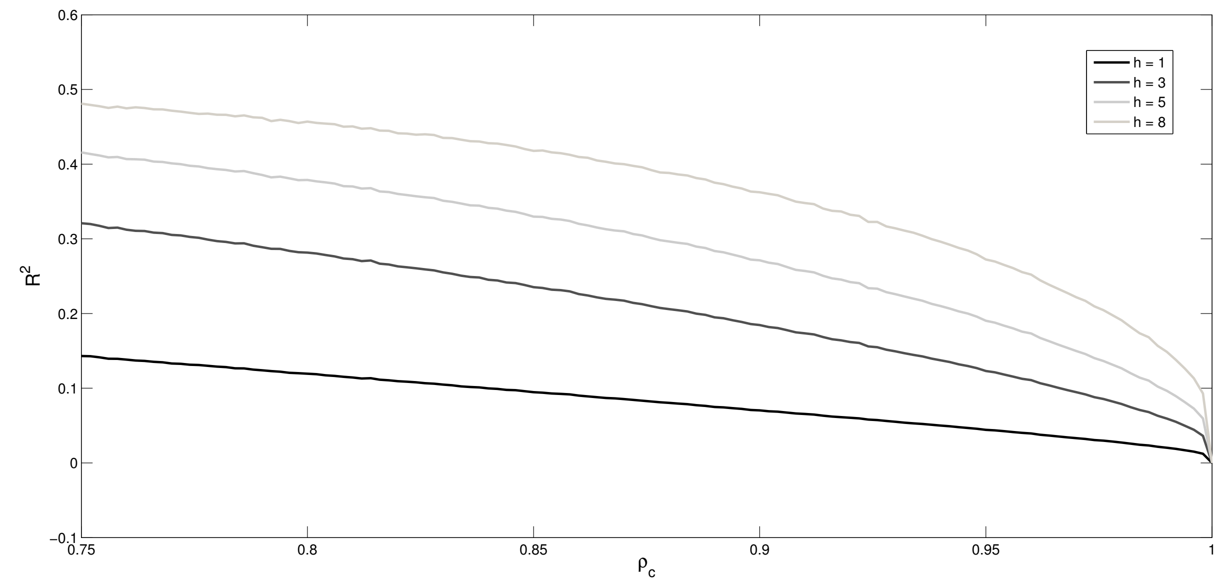
The graph shows the expected return of the aggregate consumption claim, its volatility, the Sharpe ratio, the expected risk free rate, its volatility and the equity premium as a function of the standard deviation of consumption growth $\sigma(\Delta c_t)$. *The underlying consumption process is from the garbage sample, see Table 3.1, so* $\mathrm{E}(\Delta c_t) = 0.0142$ *and* $\rho_c = 0.7661$. *The results are computed using the parameter estimates from Table 3.7 (*$\gamma = 8.24$, $\beta = 1.08$*). (*$\sigma_\nu = 0$*).*

Figure 3.17: Influence of the Autocorrelation Coefficient ρ_c on the Predictability of Stock Returns for the Post-War Dataset



The graph shows the predictability of stock returns R^2 for different forecasting horizons h as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the post-war sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.0213$ and $\sigma(\Delta c_t) = 0.0180$. The results are computed using the parameter estimates from Table 3.7 ($\gamma = 16.5$, $\beta = 1.34$). ($\sigma_\nu = 0$).

Figure 3.18: Influence of the Autocorrelation Coefficient ρ_c on the Predictability of Stock Returns for the Garbage Dataset



The graph shows the predictability of stock returns R^2 for different forecasting horizons h as a function of the autocorrelation coefficient in consumption, ρ_c. The underlying consumption process is from the garbage sample, see Table 3.1, so $\mathrm{E}(\Delta c_t) = 0.0142$ and $\sigma(\Delta c_t) = 0.0286$. The results are computed using the parameter estimates from Table 3.7 ($\gamma = 8.24$, $\beta = 1.08$). ($\sigma_\nu = 0$).

Essay 4

Computational Methods for Asset Pricing Models

Computational Methods for Asset Pricing Models with Epstein-Zin Preferences

Ole Wilms[1]
Dept. of Business Administration
University of Zurich
Moussonstrasse 15
8044 Zurich, Switzerland
ole.wilms@business.uzh.ch

November 2015

Abstract

This paper compares solution methods for solving asset pricing models with preferences of Epstein-Zin type. For this purpose, we first describe a projection-based algorithm for solving such models and compare it to the commonly used methods like log-linearization and discretization. We find that the accuracy of the commonly used methods highly depends on the model specification and approximation errors increase significantly with the persistences of the state processes. For example, for the recent calibration by Bansal, Kiku, and Yaron (2012a) of the influential Bansal and Yaron (2004) long-run risks model, using log-linearization induces errors in the log price-dividend ratio of more than 26%. In contrast, the projection method provides highly accurate approximations that are robust with regard to changes in the model parameters.

Keywords: Asset pricing; Epstein-Zin preferences; Discretization; Log-linearization; Projection methods.

[1]I am heavily indebted to my advisor Karl Schmedders as well as to Walter Pohl and Kenneth Judd for their support and guidance in this project.

4.1 Introduction

This paper compares solution methods for solving asset pricing models with recursive preferences of Epstein-Zin type (Epstein and Zin (1989) and Weil (1989)). For this purpose, we describe a projection-based algorithm for solving such models and compare it to log-linearization and discretization methods. We apply the methods to two asset pricing models. First, we consider the simple endowment economy of Tallarini (2000). The exercise serves to analyze the factors that drive the accuracy of the different approximation methods and to understand why, and in which cases, the methods fail to compute accurate solutions. Afterwards, we compare the methods for solving the long-run risk model of Bansal and Yaron (2004).

Asset pricing models with Epstein-Zin preferences generally don't have closed-form solutions. So a natural question that arises is: How do we solve such models? The commonly used method for such models is log-linearization (Bansal and Yaron (2004), Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012a), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others). While linearization methods are easy to implement and allow for approximate closed-form solutions in many cases (Eraker (2008), Eraker and Shaliastovich (2008)), they miss non-linear dynamics by construction that potentially introduce approximation errors. In this study we find that the accuracy of log-linearization strongly depends on the model specification and document large errors especially for models with highly persistent state processes. For example, for the recent calibration by Bansal, Kiku, and Yaron (2012a) of the influential Bansal and Yaron (2004) long-run risks model, using log-linearization induces errors in the volatility of the log price-dividend ratio of more than 26%.

So if the log-linearization has the potential to introduce large errors, what should we use instead? We consider two alternative families of methods. One method is to replace the continuous state space by a finite-state Markov chain and solve the discretized model. Mehra and Prescott (1985), in their original paper, use a two-state Markov chain. More general methods were introduced by Tauchen (1986) and Tauchen and Hussey (1991), and have been applied by many subsequent works, such as Guvenen (2009), Heaton and Lucas (1996), Heaton and Lucas (2000), Hansen, Heaton, and Yaron (1996), and Campbell (1993). We show that these methods are not adequate for the task at hand, and in fact can provide worse approximations than the log-linearization for the newest generation of asset pricing models, unless the number of nodes is very large. For example, in the recent calibration of the long-run risks model by Bansal, Kiku, and Yaron (2012a), using even 50 nodes for each state variable is insufficient to accurately approximate the second-order moments for asset prices, while it already takes several hours to compute solutions.

In contrast, projection methods (Judd (1992)) work very well. Projection methods, which ultimately derive from numerical methods to solve partial differential equations in physics, are conceptually more complex than log-linearization or discretization techniques. We find, however, that they work extremely well in solving general asset pricing models with Epstein-Zin preferences and that surprisingly low-dimensional approximations provide high-levels of accuracy. (See Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012) for similar success in the stochastic growth case.) The method is also robust to changes in model parameters. Accurate solutions for the two-dimensional long-run risk model can be computed in less than a second, suggesting that the approach can also be used to solve higher dimensional models.

The paper is organized as follows. Section 4.2 describes the general projection algorithm and how it can be applied to solve asset pricing models with Epstein-Zin preferences. In Section 4.3 we apply the three methods, linearization, discretization and projection, to solve the endowment economy of Tallarini (2000) and in Section 4.4 we compare the methods for solving the long-run risk model of Bansal and Yaron (2004). Section 4.5 concludes.

4.2 Model Description and Computational Approach

We consider a standard asset pricing model where the representative investor has recursive preference as in Epstein and Zin (1989) and Weil (1990). For these preferences, Epstein and Zin (1989) show that the gross return of asset i, $R_{i,t+1} = \frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}$, must satisfy the following pricing equation,

$$E_t \left[\delta^\theta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta-1} R_{i,t+1} \right] = 1 \tag{4.1}$$

where C_t is consumption, δ is the time discount factor, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$ is a parameter for the preference of the timing of risks. θ depends on the level of risk aversion γ and the intertemporal elasticity of substitution ψ. For $\theta < 1$ the agent has a preference for the early resolution of risks. For $\theta = 1$ (or equivalently $\gamma = \frac{1}{\psi}$) the utility function simplifies to the case of standard CRRA utility. The term $R_{w,t+1}$ denotes the return on a claim to aggregate consumption, $R_{w,t+1} = \frac{W_{t+1}}{W_t - C_t}$, where W_t is the (unobserved) wealth level of the agent at time t. As equation (4.1) has to hold for all assets i, it must also hold for the return of the aggregate consumption claim. So $R_{w,t+1}$ is determined by the wealth-Euler equation given by

$$E_t \left[\delta^\theta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta} \right] = 1. \tag{4.2}$$

We present an algorithm using projection methods to solve the general asset pricing model given by equations (4.1) and (4.2). Projection methods are a general-purpose tool for solving function equations. They were first introduced by physicists and engineers to solve partial differential equations, but they can be used to solve the types of fixed-point equations that arise in economics (see Judd (1992) for an introduction or Chen, Cosimano, and Himonas (2014) for a brief overview). In the following we first briefly describe the idea of projection methods and then apply it to the asset pricing model given above. Our description does not strive for maximal generality but instead is meant to simply convey the key steps of projection methods.

4.2.1 Projection Methods for Functional Equations

Projection methods are a general tool to solve functional equations of the form

$$(\mathcal{G}z)(x) = 0, \tag{4.3}$$

where the variable x resides in a (state) space $X \subset \mathbb{R}^l$, $l \geq 1$, and z is an unknown solution function with domain X, so $z : X \to \mathbb{R}^m$. The given operator $\mathcal{G}$ is a continuous mapping between two function spaces. Note that solving equation (4.3) requires finding an element z in a function space – that is, in an infinite-dimensional vector space.

The first central step of a projection method is to approximate the unknown function z on its domain X by a linear combination of basis functions. For the applications in this paper, it suffices to assume that the domain X is bounded and that the basis functions are polynomials.[2] For a set $\{\Lambda_k\}_{k\in\{0,1,\ldots,n\}}$ of chosen basis functions the approximation $\hat{z}$ of z is

$$\hat{z}(x; \boldsymbol{\alpha}) = \sum_{k=0}^{n} \alpha_k \Lambda_k(x), \tag{4.4}$$

where $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_n]$ are unknown coefficients. Replacing the function z in equation (4.3) by its approximation $\hat{z}$, we can define the residual function $\hat{F}(x; \boldsymbol{\alpha})$ as the error in the original equation,

$$\hat{F}(x; \boldsymbol{\alpha}) = (\mathcal{G}\hat{z})(x; \boldsymbol{\alpha}). \tag{4.5}$$

Instead of solving equation (4.3) for the unknown function z, we now attempt to choose coefficients $\boldsymbol{\alpha}$ to make the residual $\hat{F}(x; \boldsymbol{\alpha})$ zero. Note that instead of finding an element in an infinite-dimensional vector space we are now looking for a vector in $\mathbb{R}^{n+1}$. Obviously, this approximation step greatly simplifies the mathematical problem.

This problem is unlikely to have an exact solution, so the second central step of a projection method is to impose certain conditions on the residual function, the so-called "projection" conditions, to make the problem solvable. In other words, the purpose of the projection conditions is to establish a set of requirements that the coefficients $\boldsymbol{\alpha}$ must satisfy. For a formulation of the projection conditions, define a "weight function" (term) $w(x)$ and a set of "test" functions $\{g_k(x)\}_{k=0}^{n}$. We can then define an inner product between the residual

[2] In addition to polynomial approximations, approximations using cubic splines or B-splines are often very useful.

function $\hat{F}$ and the test function g_k,

$$\int_X \hat{F}(x;\boldsymbol{\alpha})g_k(x)w(x)dx.$$

This inner product induces a norm on the function space X. Natural restrictions for the coefficient vector α are now the projection conditions,

$$\int_X \hat{F}(x;\boldsymbol{\alpha})g_k(x)w(x)dx = 0,\ k = 0, 1, \ldots, n. \tag{4.6}$$

Observe that this system of equations imposes $n+1$ conditions on the $(n+1)$-dimensional vector α. Different projection methods vary in the choice of the weight function and the set of test functions. In this paper we describe two different projections, the collocation and the Galerkin method.

A collocation method chooses $n+1$ distinct nodes in the domain, $\{x_k\}_{k=0}^n$, and defines the test functions g_k by

$$g_k(x) = \begin{cases} 0 \text{ if } x \neq x_k \\ 1 \text{ if } x = x_k. \end{cases}$$

With a weight term $w(x) \equiv 1$, the projection conditions (4.6) simplify to

$$\hat{F}(x_k;\boldsymbol{\alpha}) = 0,\ k = 0, 1, \ldots, n. \tag{4.7}$$

Simply put, the collocation method determines the coefficients in the approximation (4.4) by solving the square system (4.7) of nonlinear equations.

The Galerkin method uses the fact that Chebyshev polynomials are orthogonal on $[-1, 1]$ with respect to the inner product using the weight function $w(x) \equiv \frac{1}{\sqrt{1-x^2}}$. Hence the Galerkin method uses the basis functions as the test functions, $g_k(x) = \Lambda_k(x)$ and the projection conditions (4.6) become

$$\int_X \hat{F}(x;\boldsymbol{\alpha})\Lambda_k(x)\frac{1}{\sqrt{1-x^2}}dx = 0,\ k = 0, 1, \ldots, n. \tag{4.8}$$

Next we show how to apply the general projection approach to solve the equilibrium pricing equations (4.1) and (4.2).

4.2.1.1 Projection Methods Applied to Asset Pricing Models

To apply a projection method to the asset pricing model, we express the equilibrium conditions as a functional equation of the type (4.3). For this purpose, we need to choose an appropriate state space and perform the usual transformation from an equilibrium described by infinite sequences (with a time index t) to the equilibrium being described by functions of some state variable(s) x on a state space X. We denote the current state of the economy by x and the subsequent state in the next period by x'. (For example in the original model by Mehra and Prescott (1985), the state x is log consumption growth and $X \subset \mathbb{R}^1$; in the model of Bansal and Yaron (2004), the state x consists of the long-run mean of consumption growth (denoted by x_t in that paper) and the variance of consumption growth (denoted by σ_t^2), so $X \subset \mathbb{R}^2$.) We assume that the probability distribution of next period's state x' conditional on the current state x is defined by a density f_x.

First note that we solve the model in two steps. In the first step, we use the projection method to solve the wealth-Euler equation (4.2) to obtain the return on wealth. Once the return on wealth is known, then, in a second step, we can solve for any asset return by applying the projection approach to equation (4.1). For the first step, write equation (4.2) in state-space representation

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + \theta r_w(x'|x)\right)\middle| x\right] = 1, \quad \forall x, \tag{4.9}$$

where lower case letters denote logs of variables and $\Delta c(x'|x) = c(x') - c(x)$. We write the model in logs, because the function we solve for is the log wealth-consumption ratio $z_w(x) = \log\left(\frac{W(x)}{C(x)}\right)$. Next, write the state-dependent log return of the aggregate consumption claim as

$$\begin{aligned} r_w(x'|x) &= \log\left(\frac{W(x')}{W(x)-C(x)}\right) = \log\left(\frac{\frac{W(x')}{C(x')}}{\frac{W(x)}{C(x)}-1} \times \frac{C(x')}{C(x)}\right) \\ &= z_w(x') - \log\left(e^{z_w(x)} - 1\right) + \Delta c(x'|x). \end{aligned} \tag{4.10}$$

Inserting the last term in equation (4.9) yields

$$E\left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + z_w(x') - \log\left(e^{z_w(x)}-1\right)\right)\right) - 1\middle| x\right] = 0, \quad \forall x. \tag{4.11}$$

Equivalently,

$$0 = \int_X \left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + z_w(x') - \log\left(e^{z_w(x)}-1\right)\right)\right) - 1\right] \mathrm{d}f_x \tag{4.12}$$

which is a functional equation of the form (4.3) and allows us to apply the projection approach.

The unknown solution function to this equilibrium condition, z_w, is an element of a function space which is an infinite-dimensional vector space. A key feature of every projection method is to approximate the solution function z_w by an element from a finite-dimensional space. Specifically, we use the approximation $\hat{z}_w(x;\boldsymbol{\alpha_w}) = \sum_{k=0}^{n} \alpha_{w,k}\Lambda_k(x)$, where $\{\Lambda_k\}_{k\in\{0,1,\ldots,n\}}$ is a set of chosen (known) basis functions and $\boldsymbol{\alpha_w} = [\alpha_{w,0}, \alpha_{w,1}, \ldots, \alpha_{w,n}]$ are unknown coefficients. Replacing the exact solution $z_w(x)$ by the approximation $\hat{z}_w(x;\boldsymbol{\alpha_w})$ leads us to the residual function $\hat{F}_w$ for the rearranged wealth-Euler equation (4.12), which is defined by

$$\hat{F}_w(x;\boldsymbol{\alpha_w}) = \int_X \left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + \hat{z}_w(x') - \log\left(e^{\hat{z}_w(x)} - 1\right)\right)\right) - 1\right] \mathrm{d}f_x. \tag{4.13}$$

We can determine values for the unknown solution coefficients $\boldsymbol{\alpha_w}$ by imposing the collocation or Galerkin projection conditions on the residual term $\hat{F}_w(x;\boldsymbol{\alpha_w})$ (see Section 4.2.1). The values for the coefficients $\boldsymbol{\alpha_w}$ determine the state-dependent wealth-consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$ which in turn leads to the (approximate) return function of the aggregate consumption claim, $\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) = \hat{z}_w(x';\boldsymbol{\alpha_w}) - \log\left(e^{\hat{z}_w(x;\boldsymbol{\alpha_w})} - 1\right) + \Delta c(x'|x)$.

With $\hat{r}_w(x'|x;\boldsymbol{\alpha_w})$ at hand, we can now develop an approach to compute the return of any asset i using equation (4.1). Analogous to the first step, we solve for the log price-dividend ratio $z_i(x) = \log\left(\frac{P(x)}{D(x)}\right)$ and rewrite the state-dependent log return of asset i as

$$\begin{aligned} r_i(x'|x) &= \log\left(\frac{P_i(x') + D_i(x')}{P_i(x)}\right) = \log\left(\frac{\frac{P_i(x')}{D_i(x')} + 1}{\frac{P_i(x)}{D_i(x)}} \times \frac{D_i(x')}{D_i(x)}\right) \\ &= \log\left(e^{z_i(x')} + 1\right) - z_i(x) + \Delta d_i(x'|x). \end{aligned} \tag{4.14}$$

Writing the Euler equation (4.1) in state-space representation and formulating it in logs yields

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta-1)r_w(x'|x) + r_i(x'|x)\right)\middle| x\right] = 1. \tag{4.15}$$

Substituting the return expressions (4.10) and (4.14) into this equations and replacing the log price-dividend ratio $z_i(x) = p_i(x) - d_i(x)$ by its approximation $\hat{z}_i(x;\boldsymbol{\alpha_i}) = \sum_{k=0}^n \alpha_{i,k}\Lambda_k(x)$ leads to the residual function

$$\begin{aligned} \hat{F}_i(x;\boldsymbol{\alpha_i}) = \int_X \Bigg[\exp\Bigg(&\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta-1)\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) \\ &+ \log\left(e^{\hat{z}_i(x';\boldsymbol{\alpha_i})} + 1\right) - \hat{z}_i(x;\boldsymbol{\alpha_i}) + \Delta d_i(x'|x)\Bigg) - 1\Bigg] \mathrm{d}f_x. \end{aligned} \tag{4.16}$$

Recall that the coefficients $\boldsymbol{\alpha_w}$ and thus the function $\hat{r}_w(x'|x;\boldsymbol{\alpha_w})$ have been computed previously. Therefore, we can now apply one of the projection conditions to solve for the unknown vector $\boldsymbol{\alpha_i}$.

In sum, we apply the projection method twice. In the first step, we approximate the log wealth-consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$ by applying the projections on the residual function of the wealth-Euler equation (4.13). Once $\boldsymbol{\alpha_w}$ is known, the projections can be applied to equation (4.16) to solve for the price-dividend ratio $\hat{z}_i(x;\boldsymbol{\alpha_i})$ of any asset i. Formally, the algorithm can be described as follows.

Algorithm *Solving Asset Pricing Models with Recursive Preferences.*

Initialization. *Define the state space* $X \subset \mathbb{R}^l$*; choose the functional forms for* $\hat{z}_w(x; \boldsymbol{\alpha_w})$ *and* $\hat{z}_i(x; \boldsymbol{\alpha_i})$ *as well as the projection method.*

Step 1. *Use the wealth-Euler equation* (4.2) *together with the approximated log wealth-consumption ratio* $\hat{z}_w(x; \boldsymbol{\alpha_w})$ *and the definition of the return equation* (4.10) *to derive the residual function for the return on wealth*

$$\hat{F}_w(x; \boldsymbol{\alpha_w}) = \int_X \left[\exp\left(\theta \left(\log\delta + (1 - \frac{1}{\psi})\Delta c(x'|x) + \hat{z}_w(x') - \log\left(e^{\hat{z}_w(x)} - 1 \right) \right) \right) - 1 \right] \mathrm{d}f_x.$$

Compute the unknown solution coefficients $\boldsymbol{\alpha_w}$ *by imposing the projections on* $\hat{F}_w(x; \boldsymbol{\alpha_w})$*.*

Step 2. *Use the solution for the wealth-consumption ratio* $\hat{z}_w(x; \boldsymbol{\alpha_w})$ *and the Euler equation* (4.1) *for asset* i *together with the approximated log price-dividend ratio* $\hat{z}_i(x; \boldsymbol{\alpha_i})$ *and the definition of the return equation* (4.14) *to derive the residual function for asset* i*,*

$$\begin{aligned} \hat{F}_i(x; \boldsymbol{\alpha_i}) = \int_X \Bigg[\exp\Bigg(& \theta \log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta - 1)\hat{r}_w(x'|x; \boldsymbol{\alpha_w}) \\ & + \log\left(e^{\hat{z}_i(x'; \boldsymbol{\alpha_i})} + 1 \right) - \hat{z}_i(x; \boldsymbol{\alpha_i}) + \Delta d_i(x'|x) \Bigg) - 1 \Bigg] \mathrm{d}f_x \end{aligned}$$

Compute the unknown solution coefficients $\boldsymbol{\alpha_i}$ *by imposing the projections on* $\hat{F}_i(x; \boldsymbol{\alpha_i})$*.*

Evaluation. *Choose a set of evaluation nodes* $\mathbb{X}^e = \{x_j^e : 1 \le j \le m^e\} \subset X$ *and compute approximation errors in the residual function of the wealth portfolio and the residual function of asset* i*. If the errors do not satisfy a predefined error bound, start over at Initialization and change the number of approximation nodes or the degree of the basis functions.*

To actually implement the algorithm, we need to specify additional algorithmic details such as the choices for basis functions and the integration technique.

4.2.1.2 Algorithmic Ingredients

In the **Initialization** step, we need to choose a set of basis functions for the polynomial approximation, a projection method and a set of nodes. To simplify the presentation, we describe the necessary choices for a one-dimensional state space approximated over an interval $X = [x_{min}, x_{max}]$. We approximate the solution functions z_w and z_i by Chebyshev polynomials (of the first kind), see Judd (1998). We obtain the Chebyshev polynomials via the recursive relationship

$$T_0(\xi) = 1, \quad T_1(\xi) = \xi, \quad T_{k+1}(\xi) = 2\xi T_k(\xi) - T_{k-1}(\xi),$$

with $T_k : [-1, 1] \to \mathbb{R}$. Since we need to approximate functions on the domain X and the Chebyshev polynomials are defined on the interval $[-1, 1]$, we need to transform the argument for the polynomials. The basis functions for the approximate solutions $\hat{z}_w(x; \boldsymbol{\alpha_w})$ and $\hat{z}_i(x; \boldsymbol{\alpha_i})$ are given by

$$\Lambda_k(x) = T_k\left(2\left(\frac{x - x_{min}}{x_{max} - x_{min}}\right) - 1\right) \tag{4.17}$$

for $k = 0, 1, \ldots, n$.

In this paper we only show the results using the collocation method but we verified the solutions using the Galerkin approach. The application of a projection method requires a set of nodes, $\mathbb{X} = \{x_j : 0 \leq j \leq m\} \subset X$; we choose the $m + 1$ zeros of the Chebyshev polynomial T_{m+1}. These points are called Chebyshev nodes,

$$\xi_j = \cos\left(\frac{2j + 1}{2m + 2}\pi\right), \qquad j = 0, 1, \ldots, m.$$

Since all Chebyshev nodes are in the interval $[-1, 1]$, we need to transform them to obtain nodes in the state space X. This transformation is

$$x_j = x_{min} + \frac{x_{max} - x_{min}}{2}(1 + \xi_j), \qquad j = 0, 1, \ldots, m.$$

For the collocation method, the number of basis functions, $n + 1$, must be identical to the number of approximation nodes, $m + 1$, and so $m = n$. In **Step 1** (and **Step 2**, if applicable), we must solve the projection conditions involving the residual function. The residual functions defined in equations (4.13) and (4.16) contain a conditional expectations operator, which also requires numerical calculations. The underlying exogenous processes in the models we consider are normally distributed, and so we apply Gauss-Hermite quadrature to calculate expectations.

The collocation approach leads to a square system of nonlinear equations, see Section 4.2.1, which can be solved with a standard nonlinear equation solver. The Galerkin projection is slightly more complex, and uses integral operators as projection conditions; these in turn can be accurately approximated by Gauss-Chebyshev quadrature.

For the **Evaluation** step we use $m^e >> m$ equally spaced evaluation nodes in X to evaluate the errors in the residual function. In particular, for asset i we compute the root mean squared errors (RMSE) and maximum absolute errors (MAE) in the residual function (4.16); these errors are

$$\text{RMSE}_i = \sqrt{\frac{1}{m^e}\sum_{j=1}^{m^e} \hat{F}_i(x_j^e|\boldsymbol{\alpha_i})^2}, \tag{4.18}$$

$$\text{MAE}_i = \max_{j=1,2,\ldots,m^e} |\hat{F}_i(x_j^e|\boldsymbol{\alpha_i})|, \tag{4.19}$$

respectively, with

$$x_j^e = x_{min} + \frac{x_{max} - x_{min}}{m^e - 1}(j-1), \qquad j = 1, \ldots, m^e. \tag{4.20}$$

4.2.2 Comparison of Three Families of Methods

In the following two sections we compare the performance of the projection methods to commonly used approaches in the literature for solving asset pricing models with recursive preferences. The two most prominent methods are discretization and linearization techniques. Specifically, we focus on the discretization methods by Tauchen (1986) and Tauchen and Hussey (1991)[3] and the log-linearized pricing kernel approach as described in Bansal and Yaron (2004). Appendix 4.A provides a brief description of the alternative solution methods for which we present results in this paper.

The common approach to solve for equilibrium dynamics is to log-linearize the model around its steady state. A discussion of log-linearization methods requires careful attention to several important differences among some well-known approaches. Standard log-linearization methods as in Judd (1996) or Collard and Juillard (2001) linearize around the deterministic steady state of the model. In a deterministic model, recursive preferences collapse to the case of CRRA preferences and hence the risk aversion has no influence (as there is no risk). But if the risk aversion has significant influence in the stochastic model, linearizing around the deterministic steady state might not be the best choice. Therefore new techniques have been developed that linearize around the risky steady state of the model (see, for example, Juillard (2011), de Groot (2013) or Meyer-Gohde (2014)).[4] Another drawback of the standard log-linearization is that the policies are independent of the volatility of the model (see Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012)). But as Bansal and Yaron (2004) point out, stochastic volatility is one of the key features of the long-run risks model and essential for asset pricing dynamics. Hence a log-linear approximation for asset pricing models with recursive preferences and stochastic volatility must account for both features, the risk-adjustment of the steady state and the effects of volatility. Bansal and Yaron (2004) use a linearization

[3] Another discretization method available for one-dimensional AR(1) processes (see e.g. Galindev and Lkhagvasuren (2010) or Kopecky and Suen (2010)) is that of Rouwenhorst (1995). Unfortunately, there is, to the best of our knowledge, no generalization of the method to dimensions higher than one.

[4] These authors define the risky steady state as the state where, in absence of shocks in the current period, the agent decides to stay at the current state while expecting shocks in the future and knowing their probability distribution.

technique based on the Campbell and Shiller (1988a) return approximation that meets these requirements which, therefore, has been used extensively for solving asset pricing models with recursive preferences (Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012a), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others).[5] One reason for its popularity is that it allows for quasi-closed form solutions for many different model specifications, for example when shocks to the economy are normal.

For the comparison of the different solution methods we report numerical solutions, error measures, and running times for two well-known asset pricing models with recursive preferences. The first asset pricing model is the endowment economy from Tallarini (2000). Log consumption is modeled as AR(1) deviations from a deterministic linear trend. The deviation from trend is the only state variable in the model and so the state space is one-dimensional. As the model is rather simple and does not account for many empirical features found in the data, we rather view it as a technical analysis to understand the strengths and weaknesses of the different solution methods, in particular with regard to changes in the preference and model parameters. (This exercise is similar in spirit to the approach in Collard and Juillard (2001), except that Collard and Juillard (2001) only consider the case of CRRA utility and hence only focus on the standard log-linearization approach around the deterministic steady state as in Judd (1996).) For the second example, we consider the long-run risk model of Bansal and Yaron (2004), which has gathered much attention recently for its ability to match many financial market characteristics. This model has a two-dimensional state space.

For both models, we first compute Euler errors in the pricing equations on the continuous state space for the projection methods and the log-linearization approach.[6] We consider a continuous state space of $\pm n_\sigma$ standard deviations around the mean of the stationary distribution for each state variable. For the computation of Euler errors, we use $m^e = 100n_\sigma$ equally spaced evaluation nodes (in each dimension) in the interval $[x_{min}, x_{max}]$ to evaluate the (absolute) error in the pricing equation. We use 8 quadrature nodes for both the Gauss-Hermite quadrature to solve the integral in the pricing equation and the Gauss-Chebyshev quadrature for the integral that arises in the Galerkin projection.[7] Regarding the efficiency of the projection approach, practical experience has shown that good initial guesses can lead to significant reductions of computation times especially for higher-order approximations. Hence we take the lower degree

[5] Another approach, proposed by Kogan and Uppal (2001) and used for example in Hansen, Heaton, Lee, and Roussanov (2007) and Hansen, Heaton, and Li (2008), is to linearize around the special case of unit elasticity of substitution $\psi = 1$ where the wealth-consumption ratio is constant. However most of the follow-up work in the long-run risk literature has focused on the log-linearization used in Bansal and Yaron (2004), which is why we concentrate on this approximation.

[6] For the discretization methods it is not straight forward how to compute errors in the Euler equation as the method only solves for the pricing functions at the discretization nodes. Errors in between the nodes could only be computed be assuming a certain interpolation procedure. As the errors would highly depend on the interpolation method used, we only provide the errors for the log-linearization and projection method.

[7] We also considered solutions with more quadrature nodes, but the results did not change significantly.

solutions as initial guesses for the higher degree approximations. For example to compute a degree-6 approximation we first compute a degree-3 approximation and take that as an initial guess. Computations times are always stated as total times including the computation time of the initial guess. All results are computed in Matlab R2014b on a 2.9 GHz Intel Core i5 with 16 GB ram and no parallelization. We use the solver 'fmincon' with the active-set algorithm and an error tolerance of 10^{-8}. (We solve nonlinear systems of equations as optimization problems with a dummy objective function and the nonlinear equations as the constraints of the optimization problem.)

In addition to the Euler errors, we also report errors in economically meaningful variables such as the mean and standard deviation of the wealth-consumption and the price-dividend ratio. To do so, we compute the "true" solution by using a very large number of nodes in the Tauchen and Hussey (1991) procedure or using a Chebyshev approximation of extremely high degree. While these calculations might take a very long time (more than two days for the Tauchen and Hussey (1991) procedure!), they allow us to evaluate numerical results for discretizations with fewer nodes, for projections with polynomials of smaller degree, and for the linearization approach.

4.3 The Endowment Economy of Tallarini (2000)

We consider the endowment economy of Tallarini (2000). Log consumption, c_t, is modeled as simple AR(1) deviations from a linear trend,

$$\begin{aligned} c_t &= \mu t + x_t \\ x_t &= \rho x_{t-1} + \sigma_\epsilon \epsilon_t, \qquad \epsilon_t \sim N(0,1), \end{aligned}$$

where μ is the average net growth rate of consumption, ρ is the degree of persistence and the state of the economy is described by the one-dimensional process x_t. In the following analysis, we focus exclusively on the pricing of the wealth portfolio (Step 1 in our solution algorithm) and make no further assumptions about dividends in the model. For the projection methods we approximate the solution over the range of $\pm n_\sigma$ standard deviations around the mean of the stationary distribution for x_t. The first two moments of this distribution are $E(x_t) = 0$ and $\sigma(x_t) = \sigma_\epsilon/\sqrt{1-\rho^2}$.

For the log-linearization approach the log wealth-consumption ratio is a linear function of the state x_t of the economy,

$$z_w(x_t) = A_{0,w} + A_{1,w} x_t.$$

Appendix 4.A describes the derivation of the unknown solution coefficients $A_{0,w}$ and $A_{1,w}$ and Appendix 4.B reports analytical expressions for the coefficients.

4.3.1 Approximation Errors in the Euler Equations

Tables 4.1 and 4.2 show Euler approximation errors for the pricing of the wealth portfolio for the two projection methods (for degrees 3, 6, and 9), and the log-linearization approach. We report Euler errors for six combinations of the preference parameters, γ and ψ. The first

Table 4.1: Euler Approximation Errors for the Log-Linearization and Projection Methods in the Model of Tallarini (2000)

	Log-Lin	**Collocation**			**Galerkin**		
		$n = 3$	$n = 6$	$n = 9$	$n = 3$	$n = 6$	$n = 9$
				$\gamma = 2, \psi = 0.5$			
$n_\sigma = 3$	4.8e-4	5.1e-6	3.0e-10	3.0e-10	5.1e-6	3.0e-10	3.0e-10
	2.4e-4	3.5e-6	1.9e-10	1.9e-10	3.5e-6	1.9e-10	1.9e-10
$n_\sigma = 10$	0.0076	1.9e-4	6.3e-8	2.9e-8	1.9e-4	6.3e-8	2.9e-8
	0.0028	1.2e-4	3.8e-8	5.4e-9	1.2e-4	3.8e-8	5.4e-9
				$\gamma = 10, \psi = 0.1$			
$n_\sigma = 3$	0.0912	0.0024	1.9e-6	1.1e-8	0.0024	3.4e-06	1.1e-8
	0.0305	0.0011	8.7e-7	7.4e-9	0.0011	1.0e-6	7.4e-9
$n_\sigma = 10$	289	0.1266	0.0018	2.4e-4	0.1266	0.0018	2.4e-4
	12.52	0.0451	0.0011	1.2e-4	0.0451	0.0011	1.2e-4
				$\gamma = 10, \psi = 1.5$			
$n_\sigma = 3$	3.0e-4	1.9e-6	7.8e-11	8.2e-11	1.9e-6	8.1e-11	8.3e-11
	1.3e-4	1.3e-6	7.3e-11	7.4e-11	1.3e-6	7.3e-11	7.4e-11
$n_\sigma = 10$	0.0035	7.3e-5	6.3e-10	9.5e-11	7.3e-5	6.4e-10	9.4e-11
	0.0015	4.8e-5	3.9e-10	7.6e-11	4.8e-5	3.9e-10	7.4e-11

The table shows Euler approximation errors for the pricing of the wealth return in the model of Tallarini (2000) for the log-linearization and the projections methods for different degrees of polynomial approximation n. Errors are reported for different sets of preference parameters γ and ψ. The first row for each pair of parameters shows the MAE and the second row the RMSE at $m_e = 100n_\sigma$ uniformly distributed evaluation nodes within n_σ standard deviations around the mean of the stationary distribution of the model.

two combinations correspond to CRRA preferences with a low ($\gamma = 2$, $\psi = 0.5$) and a high ($\gamma = 10$, $\psi = 0.1$) degree of risk aversion, respectively. The third combination is the parameter estimates from Bansal and Yaron (2004) ($\gamma = 10$ and $\psi = 1.5$). Table 4.1 reports Euler errors for these three cases. Table 4.2 depicts errors for three more cases of Epstein-Zin preferences, namely for $\gamma = 10$, $\psi = 0.5$, for $\gamma = 7.5$, $\psi = 1.5$, and for $\gamma = 2$, $\psi = 1.5$. The leftmost column indicates the size of approximating interval by providing the number n_σ of standard deviations around the mean of the stationary distribution of the model. The parameters for the consumption process are taken from Pohl, Schmedders, and Wilms (2014) and are given by

$\sigma_\epsilon = 0.0343, \mu = 0.02, \rho = 0.91.$[8] We set $\delta = 0.99$. We later vary these parameters to analyze their influence on the performance of the different solution methods.

We observe that the Euler errors for the log-linearization approach depend strongly on the evaluation range, n_σ, and the preference parameters, γ and ψ. For the standard case of CRRA utility with $\gamma = 2$ and $\psi = 0.5$ the maximum absolute approximation error is as small as 0.0076 even for 10 standard deviations around the stationary mean ($n_\sigma = 10$). The Euler errors increase dramatically for large values of risk aversion ($\gamma = 10$ and $\psi = 0.1$) with the maximum error being as large as 289 for $n_\sigma = 10$. For the parameter set of Bansal and Yaron (2004), $\gamma = 10$, $\psi = 1.5$, the approximation errors become significantly smaller, so the log-linearization appears to provide a good approximation of the model.

Table 4.2: Euler Approximation Errors for the Log-Linearization and Projection Methods in the Model of Tallarini (2000) - Second Parameter Set

	Log-Lin	**Collocation**			**Galerkin**		
		$n = 3$	$n = 6$	$n = 9$	$n = 3$	$n = 6$	$n = 9$
$\gamma = 2, \psi = 1.5$							
$n_\sigma = 3$	3.3e-5	2.2e-7	1.4e-8	1.4e-8	2.2e-7	1.4e-8	1.4e-8
	1.5e-5	1.4e-7	1.3e-8	1.3e-8	1.4e-7	1.3e-8	1.3e-8
$n_\sigma = 10$	3.9e-4	8.1e-6	1.7e-8	1.7e-8	8.1e-6	1.7e-8	1.7e-8
	1.6e-4	5.3e-6	1.3e-8	1.3e-8	5.3e-6	1.3e-8	1.3e-8
$\gamma = 10, \psi = 0.5$							
$n_\sigma = 3$	0.0052	4.5e-5	1.4e-9	1.4e-9	4.5e-5	1.4e-9	1.4e-9
	0.0022	3.1e-5	6.3e-10	5.6e-10	3.1e-5	6.3e-10	5.6e-10
$n_\sigma = 10$	0.0702	0.0017	4.7e-7	4.5e-10	0.0017	4.7e-7	4.5e-10
	0.0257	0.0011	3.3e-7	8.9e-11	0.0011	3.3e-7	8.9e-11
$\gamma = 7.5, \psi = 1.5$							
$n_\sigma = 3$	2.2e-4	1.6e-6	2.0e-7	2.0e-7	1.6e-6	2.0e-7	3.6e-13
	9.5e-5	9.4e-7	1.9e-7	1.9e-7	9.4e-7	1.9e-7	2.8e-13
$n_\sigma = 10$	0.0026	5.3e-5	2.4e-7	2.4e-7	5.3e-5	2.4e-7	3.9e-13
	0.0011	3.4e-5	1.9e-7	1.9e-7	3.4e-5	1.9e-7	2.7-13

The table shows Euler approximation errors for the pricing of the wealth return in the model of Tallarini (2000) for the log-linearization and the projections methods for different degrees of polynomial approximation n. Errors are reported for different sets of preference parameters γ and ψ. The first row for each pair of parameters shows the MAE and the second row the RMSE at $m_e = 100n_\sigma$ uniformly distributed evaluation nodes within n_σ standard deviations around the mean of the stationary distribution of the model.

[8] Parameters are estimated from the Shiller dataset on annual real consumption in the U.S. for the period from 1889-2009, see `http://www.econ.yale.edu/~shiller/data.htm`. (last accessed October 17, 2014)

For the projection methods, already for the degree-6 approximations the Euler errors are several orders of magnitude smaller than those of the log-linearization approach. Moreover, increasing the approximation degree leads to highly accurate solutions even for the larger approximation interval $n_\sigma = 10$. Both projection methods are also very robust to changes in the preference parameters. These findings are confirmed for the second parameter set (see Table 4.2). Also, both the collocation and the Galerkin method produce about the same magnitude of errors for the same degree n of the polynomial approximation.

4.3.2 Approximation Errors in the Wealth-Consumption Ratio

Next we analyze the implications of errors in the Euler equation on quantities of economic interest. For this purpose we report the relative errors (in absolute value) in the unconditional mean and the standard deviation of the log wealth-consumption ratio for the different methods. In Table 4.3 (first set of preference parameters) and Table 4.4 (second set of preference parameters) we report relative errors as well as the computation times for different sets of preference parameters and approximation degrees.

We observe that the projection methods show very low approximation errors already for the degree-3 approximations and that the results are very robust across the different preference parameters. Additionally, the degree-6 solutions provide about the same accuracy as the degree-9 approximations, which suggests that already the degree-6 approximations are able to capture most of the non-linearities in the fixed approximation interval. Note, that the degree-9 approximations might further decrease approximation errors if we increased the width of the approximation range $2n_\sigma$. The log-linearization produces relatively small approximation errors for all sets of parameters except for $\gamma = 10, \psi = 0.1$ where the errors are slightly larger.

The approximation errors of the discretization techniques are rather small for the mean of the wealth-consumption ratio but for the standard deviation the methods show difficulties and a larger number of nodes is needed to obtain accurate solutions. Comparing the different discretization methods we find that the adjustment by Floden (2007) (TH-F) slightly improves on the original method (TH) as proposed in Tauchen and Hussey (1991), particularly in the speed of convergence as the number of discretization nodes increases. Tauchen (1986)'s method often exhibits even smaller errors for the 3-node discretization but does not converge as fast as TH-F. Compared to the projection methods, the discretizations perform significantly worse while requiring about the same computation time. Log-linearization is the fastest method (usually about twice as fast as the degree-3 polynomial approximations) in this example, but it does not achieve the same accuracy as the projection methods.

Table 4.3: Relative Errors in the Unconditional Mean and Standard Deviation of the Log Wealth-Consumption Ratio in the Model of Tallarini (2000)

$\gamma = 2, \psi = 0.5$

	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	5.3e-7	6.9e-10	6.9e-10	5.3e-7	4.0e-10	3.5e-10	0.0020		
$\sigma(w_t - c_t)$	3.7e-4	4.1e-10	4.1e-10	3.7e-4	2.9e-10	2.8e-10	0.0028		
Time	0.0217	0.0334	0.0425	0.0212	0.0375	0.0549	0.0149		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	6.0e-4	3.1e-4	1.5e-4	3.9e-4	7.5e-5	1.3e-5	1.5e-4	7.4e-5	4.8e-5
$\sigma(w_t - c_t)$	0.3390	0.1326	0.0561	0.2649	0.0396	0.0059	0.2548	0.0703	4.5e-4
Time	0.0248	0.0303	0.0354	0.0277	0.0356	0.0413	0.0279	0.0337	0.0395

$\gamma = 10, \psi = 0.1$

	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	7.2e-4	4.4e-7	4.5e-9	7.2e-4	4.5e-7	4.5e-9	0.0361		
$\sigma(w_t - c_t)$	0.0106	2.6e-6	1.5e-7	0.0106	2.6e-6	1.5e-7	0.0775		
Time	0.0242	0.0321	0.0424	0.0236	0.0324	0.0525	0.0120		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	0.0418	0.0185	0.0083	0.0361	0.0064	0.0010	0.0248	0.0115	1.3e-4
$\sigma(w_t - c_t)$	0.0979	0.0011	0.0072	0.3085	0.0308	0.0016	0.4572	0.1615	0.0695
Time	0.0270	0.0297	0.0377	0.0252	0.0293	0.0376	0.0224	0.0293	0.0364

$\gamma = 10, \psi = 1.5$

	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	8.7e-8	2.4e-11	2.4e-11	8.7e-8	1.3e-11	1.2e-11	3.5e-4		
$\sigma(w_t - c_t)$	1.5e-5	1.0e-11	1.0e-11	1.5e-5	1.0e-11	5.0e-12	4.0e-4		
Time	0.0224	0.033	0.0379	0.0217	0.036	0.0494	0.0141		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	6.3e-5	4.3e-5	2.5e-5	1.3e-5	5.3e-6	1.7e-6	4.8e-4	1.4e-5	1.9e-5
$\sigma(w_t - c_t)$	0.4234	0.1902	0.0862	0.2403	0.0423	0.0074	0.1477	0.0382	0.0284
Time	0.0277	0.0294	0.0406	0.0237	0.0289	0.0350	0.0226	0.0292	0.0354

The table shows relative errors in the unconditional mean and standard deviation of the log wealth-consumption ratio as well as the computation times for different sets of preference parameters. For the projection methods we set $n_\sigma = 3$ and for the Tauchen (1986) method $m_T = 2$.

Table 4.4: Relative Errors in the Unconditional Mean and Standard Deviation of the Log Wealth-Consumption Ratio in the Model of Tallarini (2000) - Second Parameter Set

				$\gamma = 2, \psi = 1.5$					
	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	2.2e-7	2.2e-7	2.2e-7	2.2e-7	2.4e-7	2.4e-7	6.2e-5		
$\sigma(w_t - c_t)$	1.5-5	1.5-5	1.5e-5	1.5e-5	1.0e-5	9.9e-6	3.1e-5		
Time	0.0225	0.0304	0.0393	0.0222	0.0340	0.0525	0.0147		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	3.0e-5	1.4-5	6.6e-6	2.2e-5	4.0e-6	8.5e-7	3.5e-5	5.3e-6	1.3e-6
$\sigma(w_t - c_t)$	0.4235	0.1902	0.0862	0.2403	0.0423	0.0074	0.1475	0.0382	0.0284
Time	0.0246	0.0295	0.0380	0.0203	0.0255	0.0325	0.0222	0.0274	0.0361
				$\gamma = 10, \psi = 0.5$					
	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	6.4e-6	4.8e-9	4.8e-9	6.4e-6	2.8e-9	2.6e-9	6.0e-4		
$\sigma(w_t - c_t)$	3.5e-4	8.6e-9	8.6e-9	3.5e-4	1.6e-9	9.0e-10	0.0021		
Time	0.0234	0.0335	0.0419	0.0236	0.0388	0.0595	0.0150		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	0.0029	0.0017	8.6e-4	0.0015	3.5e-4	7.7e-5	0.0033	2.3e-5	4.7e-4
$\sigma(w_t - c_t)$	0.3392	0.1320	0.0550	0.2625	0.0384	0.0054	0.2203	0.0722	0.0018
Time	0.0261	0.0311	0.0395	0.0216	0.0259	0.0326	0.0244	0.0285	0.0360
				$\gamma = 7.5, \psi = 1.5$					
	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	5.7e-7	1.6e-11	1.6e-11	5.7e-7	8.7e-11	8.6e-12	6.2e-5		
$\sigma(w_t - c_t)$	1.5e-5	1.0e-11	1.0e-11	1.5e-5	1.0e-11	5.0e-12	1.7e-5		
Time	0.0229	0.0336	0.0410	0.0225	0.0386	0.0567	0.0135		
	TH			**TH-F**			**Tauchen**		
	3	6	9	3	6	9	3	6	9
$E(w_t - c_t)$	3.4e-5	2.6e-5	1.5e-5	2.3e-6	2.9e-6	5.4e-7	3.43e-4	9.6e-6	1.2e-5
$\sigma(w_t - c_t)$	0.4235	0.1902	0.0862	0.2403	0.0423	0.0074	0.1475	0.0382	0.0284
Time	0.0248	0.0295	0.0347	0.0208	0.0260	0.0305	0.0236	0.0291	0.0356

The table shows relative errors in the unconditional mean and standard deviation of the wealth-consumption ratio as well as the computation times for different sets of preference parameters. For the projection methods we set $n_\sigma = 3$ and for the Tauchen (1986) method $m_T = 2$.

We briefly summarize the performance of the three families of numerical solution methods for the endowment economy of Tallarini (2000). The two projection methods deliver Euler approximation errors as well as relative errors for the first two moments of the wealth-consumption ratio that are typically several orders of magnitude smaller than the corresponding values for the discretization and log-linearization methods. However, this considerable outperformance does not appear to be relevant from the viewpoint of economics.

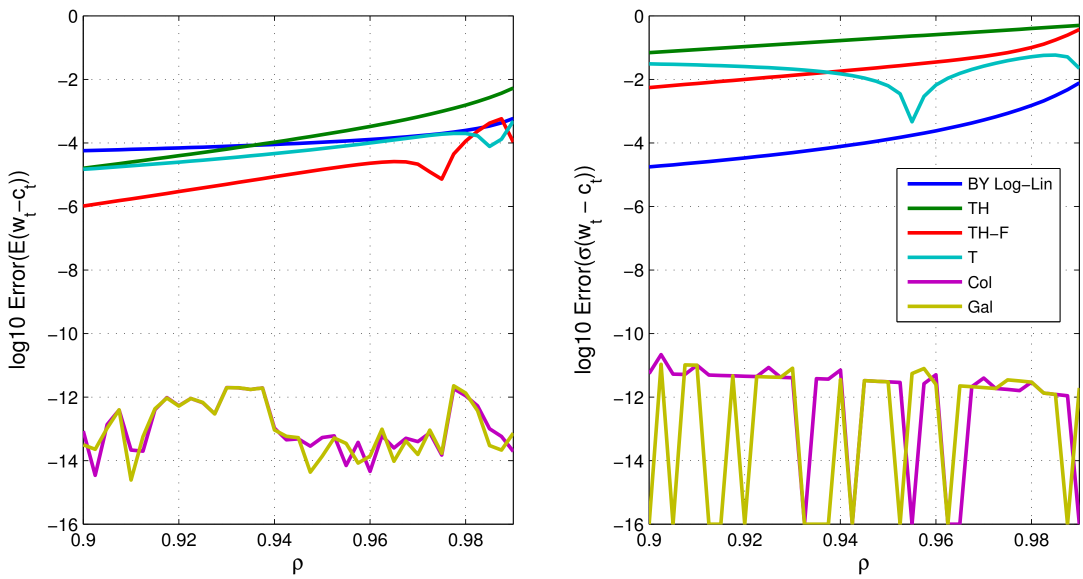
The log-linearization method has the largest relative errors for the parametrization $\gamma = 10, \psi = 0.1$, namely 3.61% for the average wealth-consumption ratio and 7.75% for its standard deviation. (The largest errors for the 9-node Tauchen procedure are smaller.) However, while errors of this size may be annoying, they hardly matter for a qualitative interpretation of an economic model. In fact, errors of this magnitude may not even matter in a quantitative economic analysis. To put it bluntly, the great numerical advantage of the projection methods over the other two families of methods does not bear relevance to the economic analysis of the model. In Section 4.4 we demonstrate that such a conclusion in favor of the simpler but inferior solution methods is not correct in general, particularly for more complex models.

In the next step of the analysis, we evaluate the robustness of the methods with regard to changes in the underlying parameter estimates. The exercise serves to analyze the factors that drive the accuracy of the different approximation methods and to understand, why and in which cases the methods fail to compute accurate solutions. In particular it delivers interesting insights regarding suitable solution methods for the long-run risks model in Section 4.4. Therefore we focus on the preference parameters $\gamma = 10$ and $\psi = 1.5$ which are the parameters used in the long-run risks model of Bansal and Yaron (2004).

4.3.3 Robustness with Regard to Changes in the Input Parameters

In this subsection we evaluate the performance of the solution methods with regard to their sensitivity to changes in the model parameters. Figure 4.1 shows the approximation errors as in Table 4.3 for different values of ρ and Figure 4.2 shows the corresponding errors for variations in σ_ϵ. We observe that the performance of the log-linearization and the discretization methods depend on the model parameters. The approximation error of the log-linearization increases strongly with the serial correlation of the underlying process and also with the volatility. (Note the log10 scale.) This result is especially interesting, as the long-run risk model considered in the next section, relies on highly persistent processes. The approximation errors of Tauchen and Hussey (1991) also increase with the serial correlation ρ. This is a well documented fact in the literature (see Floden (2007)). The method of Floden (2007) performs better, but also shows severe difficulties in approximating second-order moments. Tauchen's method is more robust with regard to changes in ρ. The errors for the projection methods on the other hand are difficult to distinguish from zero with maximum errors in the order of 10^{-10} and prove to be very robust with regard to changes in both parameters.[9]

Figure 4.1: Relative Errors in the Pricing of the Mean and Standard Deviation of the Log Wealth-Consumption Ratio as a Function of ρ in the Model of Tallarini (2000)



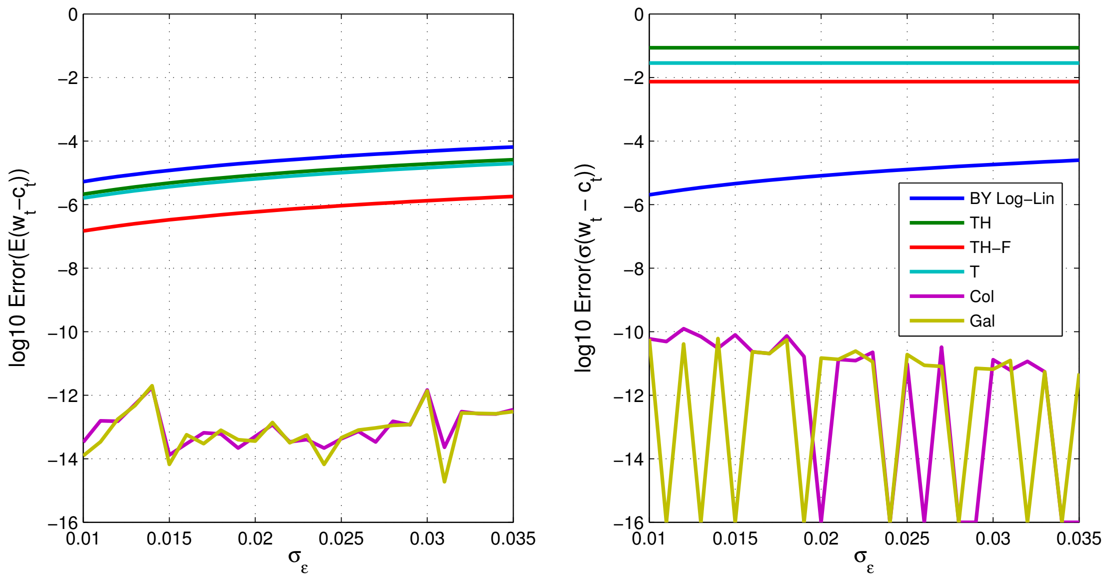
The graph shows relative errors in the unconditional mean (left panel) and standard deviation (right panel) of the wealth-consumption ratio for different values of ρ in log10 scale. For the projection methods a degree-9 approximation is used with $n_\sigma = 3$. For the discretizations 9 nodes are used and $m_T = 2$.

The results show that while all the methods can provide more or less accurate solutions for the simple endowment economy considered in this section, the accuracy of the log-linearization and the discretization methods depends strongly on the parametrization of the underlying process. In particular, we find that high serial correlation or large volatilities can significantly increase approximation errors.

In the following section, we analyze the performance of the different methods for solving the long-run risks model of Bansal and Yaron (2004). This model features highly persistent processes not only for the long run growth component but also for the stochastic volatility. Therefore, in light of our results in this subsection, we may expect that methods such as the log-linearization or the discretization approach might induce large approximation errors for models with such features.

[9]The bouncing around of the errors for the projection methods are due to getting close to machine precision. Note that the errors are in the order of 10^{-14} and hence are basically zero from a computational point of view.

Figure 4.2: Relative Errors in the Pricing of the Mean and Standard Deviation of the Log Wealth-Consumption Ratio as a Function of σ_ϵ in the Model of Tallarini (2000)



The graph shows relative errors in the unconditional mean (left panel) and standard deviation (right panel) of the wealth-consumption ratio for different values of σ_ϵ in log10 scale. For the projection methods a degree-9 approximation is used with $n_\sigma = 3$. For the discretizations 9 nodes are used and $m_T = 2$.

4.4 The Long-Run Risks Model of Bansal and Yaron (2004)

As the second application we consider the long-run risks model of Bansal and Yaron (2004). The main innovation in the model is that growth rates feature random but highly persistent long-run shocks. Additionally the conditional variance of the growth rates is itself stochastic. So the model has two state variables, the long-run component, x_t, and the variance level, σ_t^2. The model has emerged as a workhorse asset pricing model, and there have been many variations and extensions that use the same log-linear approximation as in Bansal and Yaron (2004). For example Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013) add a third highly persistent process to model inflation, Bansal, Kiku, and Yaron (2010) add a third state variable to model business cycle risk, and Kaltenbrunner and Lochstoer (2010) analyze how long-run risk arises in a production economy.

In this paper we focus on the original model from Bansal and Yaron (2004) and show that even the standard model has highly nonlinear policy functions that can introduce large approximation errors depending on the calibration of the model. Hence when using log-linear approximations for further extensions of the model, particular attention should be paid to the accuracy of the solution. Bansal and Yaron (2004) specify the original model as follows.

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
x_{t+1} &= \rho x_t + \phi_e \sigma_t e_{t+1} \\
\sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \sigma_\omega \omega_{t+1} \\
\Delta d_{t+1} &= \mu_d + \Phi x_t + \phi\sigma_t u_{t+1} + \pi\sigma_t\eta_{t+1} \\
& \quad \eta_{t+1}, e_{t+1}, \omega_{t+1}, u_{t+1} \sim i.i.d.\ N(0,1).
\end{aligned}$$

We consider two sets of parameter values for the model. The original set used in Bansal and Yaron (2004) and the more recent set in Bansal, Kiku, and Yaron (2012a). The parameters are calibrated to match annual financial market characteristics for the period from 1930–2008 while the representative agent has a monthly decision interval. Table 4.5 lists the two sets of parameter estimates. The two parameter sets mainly differ with respect to the persistences

Table 4.5: Parameters for the Long-Run Risks Model

	μ_c	ρ	ϕ_e	$\bar{\sigma}$	ν	σ_ω	μ_d	Φ	ϕ	π	γ	ψ	δ
BKY (2012a)	1.5e-3	0.975	0.038	7.2e-3	0.999	2.8e-6	1.5e-3	2.5	5.96	2.6	10	1.5	0.9989
BY (2004)	1.5e-3	0.979	0.044	7.8e-3	0.987	2.3e-6	1.5e-3	3.0	4.5	0	10	1.5	0.998

The table shows the parameters estimates of the long-run risks model as reported in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a).

of the state processes. While the original calibration uses a persistence of $\nu = 0.987$ for the stochastic volatility process, the new calibration of Bansal, Kiku, and Yaron (2012a) uses a persistence of $\nu = 0.999$ to increase the influence of the volatility channel. Also Bansal, Kiku, and Yaron (2012a) assume that shocks to consumption influence dividends to model the correlation of the two processes reported in the data ($\pi = 2.6$ in the new calibration compared to $\pi = 0$ in the original calibration). Therefore, in this paper we focus particularly on the Bansal, Kiku, and Yaron (2012a) calibration, as it displays more consistency with key statistics of financial markets data (than the original 2004 calibration).

We solve the model for the return of the wealth portfolio, z_w, the market portfolio, z_m, and the risk-free rate, z_{rf}. As in the previous section, we first examine Euler errors for the continuous methods and evaluate all methods with respect to their ability to compute unconditional moments of the model variables. For the approximation interval of the projection methods we choose the interval to be slightly larger than the maximum observation range of the long simulations. The values are given by $x_{min} = -0.013$, $x_{max} = 0.013$, $\sigma^2_{min} = 0$ and $\sigma^2_{max} = 0.00038$.[10] For the collocation method we use the full tensor product of one-dimensional basis functions, which allows us to use Chebyshev nodes in each dimension and still maintain an exactly identified system of equations–that is, $(n+1)^2$ unknown solution coefficients and $(m+1)^2$

[10] We also tried larger intervals which didn't significantly change the results.

approximation nodes with $n = m$. For the Galerkin method we choose instead the set of complete polynomials, which are the products of one-variable polynomials such that the total degree is at most $n+1$. This choice reduces the number of unknown solution coefficients from $(n+1)^2$ to $(n+1)^2/2 + (n+1)/2$ and thus lowers the computational costs without much loss of approximation quality.

For the log-linearization approach, the log wealth-consumption ratio z_w, the log price-dividend ratio z_m, and the log risk-free rate are linear functions of the state variables,

$$\begin{aligned} z_w(x_t, \sigma_t^2) &= A_{0,w} + A_{1,w} x_t + A_{2,w} \sigma_t^2 \\ z_m(x_t, \sigma_t^2) &= A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2 \\ z_{rf}(x_t, \sigma_t^2) &= A_{0,rf} + A_{1,rf} x_t + A_{2,rf} \sigma_t^2. \end{aligned}$$

Appendix 4.A describes the derivation of the unknown solution coefficients and Appendix 4.B reports analytical expressions for all nine coefficients.

4.4.1 Approximation Errors in the Euler Equations

Table 4.6 shows Euler errors for the wealth, market, and risk-free return. The first row of each entry shows the maximum absolute error and the second row the root mean squared error. We observe that already the degree-3 polynomial approximation delivers errors that are usually about 2 orders of magnitude smaller than those of the linear approximation. The collocation performs slightly better than Galerkin projection in most cases, which might be driven by the fact that we use complete polynomials for the Galerkin projection and tensor products for the collocation. As we document below, using complete polynomials can lead to significant gains in computation time. In general also the errors for the log-linearization are rather small with maximum errors of 0.0168 for the calibration of Bansal, Kiku, and Yaron (2012a) and 0.0119 for the calibration of Bansal and Yaron (2004). In the next section we analyze if this finding also holds true for economically relevant variables like errors in the moments of the wealth-consumption and price-dividend ratio.

Table 4.6: Euler Approximation Errors for the Log-Linearization and Projection Methods in the Long-Run Risks Model

	BY Log-Lin	**Collocation**			**Galerkin**		
Calibration BKY (2012a)							
		$n = 3$	$n = 6$	$n = 9$	$n = 3$	$n = 6$	$n = 9$
Wealth-Euler	0.0160	3.9e-5	5.8e-7	5.8e-7	1.5e-4	3.7e-7	2.0e-11
	0.0044	1.6e-5	1.3e-7	1.3e-7	3.9e-5	7.3e-8	3.0e-12
Market-Euler	0.0144	1.9e-4	4.0e-7	4.0e-7	3.3e-4	2.4e-6	2.1e-8
	0.0022	6.6e-5	1.5e-7	1.5e-7	9.5e-5	5.1e-7	3.5e-9
Risk-Free-Euler	0.0168	4.1e-5	7.6e-9	7.6e-9	1.5e-4	3.5e-8	1.2e-11
	0.0046	1.7e-5	4.5e-9	4.5e-9	4.0e-5	8.7e-9	2.1e-12
Calibration BY (2004)							
		$n = 3$	$n = 6$	$n = 9$	$n = 3$	$n = 6$	$n = 9$
Wealth-Euler	0.0018	1.2e-5	1.3e-8	1.3e-8	4.7e-5	3.3e-8	3.2e-8
	4.3e-4	6.3e-6	1.2e-9	1.2e-9	1.0e-5	2.7e-9	2.7e-9
Market-Euler	0.0119	2.3e-4	3.8e-7	1.0e-9	5.0e-4	3.0e-6	2.6e-8
	0.0023	1.3e-4	2.1e-7	2.1e-7	1.5e-4	5.2e-7	2.5e-9
Risk-Free-Euler	0.0019	1.2e-5	2.2e-11	3.9e-12	4.9e-5	1.2e-9	2.9e-12
	4.5e-4	6.6e-6	8.9e-12	8.9e-12	1.1e-5	1.1e-10	1.1e-12

The table shows Euler approximation errors for the pricing of the wealth return, the market return and the risk-free rate for the log-linearization and the projections methods for different degrees of polynomial approximation n. Errors are reported for the parameter specifications in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a), respectively. The first row for each Euler equation shows the maximum absolute error and the second row the root mean squared error of 200 uniformly distributed evaluation nodes in each dimension.

4.4.2 Approximation Errors in the Wealth-Consumption and Price-Dividend Ratio

Table 4.7 shows relative errors in the unconditional mean and standard deviation of the log wealth-consumption and log price-dividend ratio, as well as computation times. We find that the linearization does a reasonably good job for the parameters in Bansal and Yaron (2004) with a maximum error of 1.53% in the volatility of the price-dividend ratio. For the parameter set of Bansal, Kiku, and Yaron (2012a) the results are considerably worse. The linearization has particular difficulties in approximating second order moments, with a maximum error of 26.9% in the volatility of the price-dividend ratio.

Table 4.7: Relative Errors in the Unconditional Mean and Standard Deviation of the Log Wealth-Consumption and Log Price-Dividend Ratio in the Long-Run Risks Model

Calibration BKY (2012a)									
	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	1.5e-5	2.0e-8	2.0e-8	3.2e-6	2.1e-8	2.0e-8	0.0105		
$\sigma(w_t - c_t)$	4.6e-4	4.9e-7	4.9e-7	9.1e-4	1.5e-6	1.1e-6	0.1225		
$E(p_t - d_t)$	0.0024	2.9e-8	2.9e-8	4.4e-4	5.4e-7	4.2e-7	0.0315		
$\sigma(p_t - d_t)$	0.0035	9.4e-6	9.4e-6	0.0335	1.5e-5	1.4e-6	0.2690		
Time	1.01	4.76	12.14	0.59	3.18	9.99	0.023		
	TH			**TH-F**			**Tauchen**		
	10	30	50	10	30	50	10	30	50
$E(w_t - c_t)$	0.0794	0.0608	0.0533	0.0556	0.0432	0.0203	0.0378	0.0121	0.0056
$\sigma(w_t - c_t)$	0.9521	0.8627	0.8240	> 1	> 1	0.7682	> 1	0.7672	0.0999
$E(p_t - d_t)$	0.2533	0.1223	0.0877	0.1171	0.0816	0.0334	0.0773	0.0249	0.0060
$\sigma(p_t - d_t)$	0.8881	0.6981	0.6267	> 1	> 1	0.6992	> 1	0.6123	0.1312
Time	2.35	228.8	2944	1.67	354.7	4664	2.16	293.5	3697
Calibration BY (2004)									
	Collocation			**Galerkin**			**BY Log-Lin**		
	3	6	9	3	6	9			
$E(w_t - c_t)$	1.1e-5	7.4e-12	7.4e-12	2.1e-5	3.9e-9	3.9e-9	1.7e-4		
$\sigma(w_t - c_t)$	1.9e-4	8.4e-10	8.4e-10	2.9e-4	1.1e-8	1.1e-8	0.0013		
$E(p_t - d_t)$	0.0028	7.0e-6	7.0e-6	0.0017	2.2e-5	4.0e-8	0.0054		
$\sigma(p_t - d_t)$	0.0146	2.3e-5	2.3e-5	0.0208	2.8e-5	4.5e-8	0.0153		
Time	0.67	4.27	12.18	0.53	2.52	9.33	0.023		
	TH			**TH-F**			**Tauchen**		
	10	30	50	10	30	50	10	30	50
$E(w_t - c_t)$	0.0538	0.0238	0.0123	0.0015	2.8e-4	9.8e-5	0.0318	0.0019	0.0038
$\sigma(w_t - c_t)$	0.7739	0.3227	0.1634	0.0758	0.0123	0.0105	0.5886	0.0386	0.0560
$E(p_t - d_t)$	0.6509	0.1085	0.0518	0.0164	3.9e-4	5.6e-4	0.0635	0.0104	0.0169
$\sigma(p_t - d_t)$	0.7427	0.2968	0.15472	0.0074	0.0063	0.0033	0.2479	0.0512	0.0611
Time	2.87	215.8	2867	1.44	285.0	3792	1.87	234.9	2893

The table shows relative errors in the unconditional mean and standard deviation of the monthly log wealth-consumption and log price-dividend ratio as well as the computation times. Errors are reported for the parameter specifications in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a), respectively.

As in the previous section for the model with a one-dimensional state space, we find that the degree-6 solutions provide about the same accuracy as the degree-9 approximations; this finding suggests that already the degree-6 approximations are able to capture most of the non-linearities in the fixed approximation interval. Tauchen and Hussey (1991)'s method is not able to produce reliable results using a 10-node discretization and shows very slow convergence properties, particularly for the calibration by Bansal, Kiku, and Yaron (2012a) with its large value for the persistence parameter ν. Again we find that all discretization methods show particular difficulties in approximating the second order dynamics. The computation time of the discretization methods increases dramatically with the number of discretization nodes. For example, for 10 nodes the Tauchen and Hussey (1991) method takes about 3 seconds to compute the optimal solution, while it takes about 50 minutes to compute the 50-node approximation. The projection methods take less than two seconds to compute the degree-3 and less than five seconds for the degree-6 approximations, which already provide highly accurate results. In addition the Galerkin method is a bit faster than the collocation approach.

In sum, using log-linearized approximations or discretization techniques to solve long-run risk models can imply large errors, while projection methods provide highly accurate and robust solutions. Contrary to our experience with the endowment economy of Tallarini (2000), the large numerical errors of the log-linearization method have significant economic implications and substantially distort the relevant economic results. For example, the error using log-linearization to compute the volatility of the log price-dividend ratio is larger than 26%. Future research should analyze the influence of the errors on other economically relevant equilibrium outcomes like the annualized equity premium return volatilities or predictability patterns. Also it would be interesting how adding different state processes influences approximation errors. However, to be on the safe side, researchers should rather use more sophisticated methods, like the projection approach described in this paper, to minimize approximation errors beforehand.

4.5 Conclusion

This paper compares solution methods to solve asset pricing models with preferences of Epstein-Zin type. Projection methods, a general tool for solving functional equations, are well-suited for the approximation of nonlinear pricing functions. We have found that the projection methods constructed in this paper outperform commonly used methods such as discretization and log-linearization in terms of efficiency and accuracy. These improvements become particularly significant for the latest generation of asset pricing models, such as the influential long-run risks model. The increasing complexity of these asset pricing models requires numerical solution methods that are robust to changes in model specification.

We have shown that while the log-linearization provides a fast and easy solution method, its accuracy depends highly on the model specification. In the most recent calibration of the long-run risks model (see Bansal, Kiku, and Yaron (2012a)), the approximation errors in the volatility of the log-price dividend ratio exceed 26%. While discretization methods may, at least in theory, guarantee convergence of the approximate solution to the true equilibria, they are inefficient and highly dependent on the choice of parameters. They have severe difficulties in the presence of highly persistent consumption processes and hence require a large number of discretization nodes. But since computation times increase dramatically with the number of nodes, particularly in higher dimensions, the discretization methods are all but impractical.

The projection method presented in this paper proves to be highly accurate and the performance depends neither on the choice of preference parameters nor on the specification of the underlying consumption processes. Already the degree-3 approximations yield highly accurate solutions in the asset pricing models under consideration while they take only slightly longer than the log-linearization approach. The degree-6 approximations provide errors that are several magnitudes smaller than those of the other methods. In one dimension the difference between the Galerkin and the collocation projection is only marginal, but the Galerkin method proves to be more efficient for higher dimensions since it can be used together with complete polynomials instead of tensor products which are not easily implemented for collocation.

The results of this paper suggest that the methods, that have been used in the past, are not suitable for solving modern asset pricing models while the projection method presented in this paper provides an efficient and robust alternative.

4.A Alternative Solution Methods

In this Appendix we provide a brief description of the alternative solution methods, namely the discretization methods by Tauchen (1986) and Tauchen and Hussey (1991) and the log-linearization approach as described in Bansal and Yaron (2004), considered in this paper. The description does not strive for maximal generality but instead is meant to provide the reader with relevant information regarding the key steps and differences of the several methods. Similar to the projection algorithm, these three methods have to be conducted in two steps by first solving for the return on wealth, and then solving for the return of any individual asset.

4.A.1 Discretization

The idea of discretization methods is to discretize the continuous state space by a finite number of discretization nodes and to design a Markov transition matrix for a Markov chain on the set of nodes. Put differently, these methods replace the continuous state space and conditional density functions by a discrete state space and transition probabilities, respectively. With the nodes and the Markov transition probabilities at hand, the pricing equation (4.15) becomes a square system of nonlinear equations. This nonlinear system has as many equations as nodes; the unknown variables are the log price-dividend ratio at each node. We can then solve this system with a standard nonlinear equation solver.

Discretization methods differ in how they choose the discretization nodes and transition probabilities. For demonstration purposes, we consider the simple case of the discretization of an AR(1) process that is given by

$$x_{t+1} = (1-\rho)\mu + \rho x_t + \epsilon_{t+1}, \quad \epsilon_t \sim N(0, \sigma_e^2), \tag{4.21}$$

with persistence $|\rho| < 1$ and the unconditional mean μ. The unconditional volatility of the process is given by $\sigma_x = \sigma_e / \sqrt{1-\rho^2}$.

Tauchen (1986)'s method

Tauchen (1986) assumes a set of equally spaced nodes $X_{n_T} = \{x_1, \ldots, x_{n_T}\}$ for the discrete state space with $x_1 = \mu - m_T \sigma_y$ and $x_{n_T} = \mu + m_T \sigma_y$. The factor m_T is a positive real number and determines the range of the state space. (To the best of our knowledge there is no optimal rule for choosing m_T even though its value strongly influences the approximation results.) Denote the step size between two adjacent grid points by $h = x_i - x_{i-1}$. Then the elements π_{ij} of the $(n_T \times n_T)$-transition probability matrix π are defined by

$$\pi_{ij} = \begin{cases} \Phi\left(\frac{x_j + h/2 - (1-\rho)\mu - \rho x_i}{\sigma_e}\right) & \text{for} \quad j = 1, \\ \Phi\left(\frac{x_j + h/2 - (1-\rho)\mu - \rho x_i}{\sigma_e}\right) - \Phi\left(\frac{x_j - h/2 - (1-\rho)\mu - \rho x_i}{\sigma_e}\right) & \text{for} \quad 1 < j < n_T, \\ 1 - \Phi\left(\frac{x_j - h/2 - (1-\rho)\mu - \rho x_i}{\sigma_e}\right) & \text{for} \quad j = n_T. \end{cases}$$

Tauchen and Hussey (1991)'s method and the extension by Floden (2007)

Tauchen and Hussey (1991)'s method is based on Gauss-Hermite quadrature. Let ξ_i and ω_i, $i = 1, \ldots, n_{TH}$, be the Gauss-Hermite nodes and weights on the interval $[-\infty, +\infty]$, respectively. The approximation nodes are then given by $x_i = \mu + \sqrt{2}\sigma_e \xi_i$ and the entries π_{ij} of the transition probability matrix π can be computed by

$$\pi_{ij} = \frac{\hat{\omega}_{ij}}{\sum_{j=1}^{N} \hat{\omega}_{ij}} \tag{4.22}$$

with

$$\hat{\omega}_{ij} = \pi^{-0.5} \omega_j \frac{f(x_j|x_i)}{f(x_j|\mu)} \tag{4.23}$$

where $f(\cdot|x_i)$ is the density function of $N((1-\rho)\mu + \rho x_i, \sigma^2)$. Tauchen and Hussey (1991) simply choose $\sigma = \sigma_e$. Floden (2007) chooses $\sigma = a\sigma_e + (1-a)\sigma_x$ with $a = 0.5 + 0.25\rho$, so σ is a weighted average of σ_x and σ_e. He claims that his approach performs significantly better than the original Tauchen and Hussey (1991) method, particularly for highly persistent processes.

4.A.2 Log-Linearization as in Bansal and Yaron (2004)

Here we provide a short sketch of the linearization method that is sufficient to understand the key steps of the linearization and how it can be applied to the asset pricing models considered in this paper. For a general description of the method see Eraker (2008) and Eraker and Shaliastovich (2008). Assume that the log price-dividend ratio of asset i, $z_{i,t}$ is a linear function of the state variables

$$z_{i,t} = A_{0,i} + A_i y_t \tag{4.24}$$

where $y_t \in \mathbb{R}^l$ is the state vector describing the economy and $A_{0,i} \in \mathbb{R}^1$ and $A_i \in \mathbb{R}^l$ are the unknown linearization coefficients. The log return of the asset i, $r_{i,t+1}$ is then defined as

$$r_{i,t+1} = \log\left(e^{z_{i,t+1}} + 1\right) - z_{i,t} + \Delta d_{i,t+1} \tag{4.25}$$

where $\Delta d_{i,t+1}$ is the log growth rate of dividends.

Making use of the Campbell and Shiller (1988a) return approximation one gets

$$r_{i,t+1} \approx \kappa_{i,0} + \kappa_{i,1} z_{i,t+1} - z_{i,t} + \Delta d_{i,t+1} \tag{4.26}$$

with the linearizing constants

$$\kappa_{i,1} = \frac{e^{\bar{z}_i}}{1 + e^{\bar{z}_i}} \tag{4.27}$$

$$\kappa_{i,0} = -\log\left((1 - \kappa_{i,1})^{1-\kappa_{i,1}} \kappa_{i,1}^{\kappa_{i,1}}\right) \tag{4.28}$$

that only depend on the model implied mean price-dividend ratio $\bar{z}_i = A_{0,i} + A_i E(y_t)$. Plugging the return approximation for the return on wealth (4.26) into the equilibrium condition (4.2) yields

$$E_t\left[e^{\theta \log \delta + (\theta - \frac{\theta}{\psi})\Delta c_{t+1} + \theta(\kappa_{w,0} + \kappa_{w,1} z_{w,t+1} - z_{w,t})}\right] = 1. \tag{4.29}$$

The equilibrium condition now only depends on the state of the economy and the linearization coefficients $A_{0,i}$ and A_i. As the equilibrium equation has to hold for any realization of the state of the economy, one can collect the terms for each state to obtain a square system of $l + 1$ equations. Once we have solved for the return on wealth one can apply the linearization approach to the general pricing equation (4.1) to solve for the log price-dividend ratio of any asset i. For certain state processes the expectation can be evaluated analytically, as for example for processes with normal innovations as in Bansal and Yaron (2004) or Bollerslev, Tauchen, and Zhou (2009). This allows for quasi closed-form solutions for the linearization coefficients that only depend on the linearization constants $\kappa_{i,0}$ and $\kappa_{i,1}$. Eraker (2008), Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011) show how to generalize the approach to include general affine processes and jumps.

4.B Coefficients for the Log-Linearization

4.B.1 Log-Linearization Coefficients for the Endowment Economy of Tallarini (2000)

Parameters of the BY Log-Linearization as derived in Appendix 4.A for the Endowment Economy of Tallarini (2000) considered in Section 4.3:

$$A_{0,w} = \frac{\log \delta + (1 - \frac{1}{\psi})\mu + 0.5\theta((1 - \frac{1}{\psi}) + A_1)^2 \sigma_\epsilon^2 + \kappa_{w,0}}{\kappa_{w,1} - \rho}$$

$$A_{1,w} = \frac{(1 - \frac{1}{\psi})(\rho - 1)}{1 - \kappa_{w,1}\rho}$$

4.B.2 Log-Linearization Coefficients for the Long-Run Risks Model

Parameters of the BY Log-Linearization as derived in Appendix 4.A for the long-run risks model of Bansal and Yaron (2004) considered in Section 4.4:

$$
\begin{aligned}
A_{0,w} &= \frac{\log\delta + (1-\frac{1}{\psi})\mu + A_{2,w}\bar{\sigma}^2(1-\nu) + \kappa_{0,w} + 0.5\theta(A_{2,w}\sigma_w)^2}{\kappa_{1,w}-1} \\
A_{1,w} &= \frac{(1-\frac{1}{\psi})}{\kappa_{1,w}-\rho} \\
A_{2,w} &= \frac{0.5\theta\left((1-\frac{1}{\psi})^2 + (A_{1,w}\phi_e)^2)\right)}{\kappa_{1,w}-\nu} \\
A_{0,m} &= \Big[\theta\log\delta - \gamma\mu + \mu_d + (\theta-1)(\kappa_{0,w} + A_{0.w}(1-\kappa_{1,w})) + (\theta-1)A_{2,w}\bar{\sigma}^2(1-\nu) \\
&+ \kappa_{0,m} + \kappa_{1,m}A_{2,m}\bar{\sigma}^2(1-\nu) + 0.5\left((\theta-1)A_{2,w}\sigma_\omega + \kappa_{1,m}A_{2,m}\sigma_\omega\right)^2\Big]/(1-\kappa_{1,m}) \\
A_{1,m} &= \frac{(\Phi-\frac{1}{\psi})}{1-\kappa_{1,m}\rho} \\
A_{2,m} &= \frac{0.5((\pi-\gamma)^2+\phi^2) + 0.5((\theta-1)A_{1,w}\phi_e + \kappa_{1,m}A_{1,m}\phi_e)^2 + (\theta-1)A_{2,w}(\nu-\kappa_{1,w})}{1-\kappa_{1,m}\nu} \\
A_{0,rf} &= \theta\log\delta - \gamma\mu + (\theta-1)(\kappa_{0,w} + A_{0,w}(1-\kappa_{1,w}) + A_{2,w}\bar{\sigma}^2(1-\nu)) + 0.5((\theta-1)A_{2,w}\sigma_\omega)^2 \\
A_{1,rf} &= -\gamma + (\theta-1)A_{1,w}(\rho-\kappa_{1,w}) \\
A_{2,rf} &= 0.5(\gamma^2 + ((\theta-1)A_{1,w}\phi_e)^2) + (\theta-1)A_{2,w}(\nu-\kappa_{1,w})
\end{aligned}
$$

Part III

Bibliography and Curriculum Vitae

Bibliography

Abel, A. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 80, 38–42.

Alchian, A. A. (1950): "Uncertainty, Evolution, and Economic Theory," *Journal of Political Economy*, 58, 211–221.

Alvarez, F. and U. J. Jermann (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73, 1977–2016.

Andreou, E. and A. Spanos (2003): "Statistical Adequacy and the Testing of Trend Versus Difference Stationarity," *Econometric Review*, 22, 217–237.

Atkinson, K. E. (1992): "A Survey of Numerical Methods for Solving Nonlinear Integral Equations," *Journal of Integral Equations and Applications*, 4, 15–46.

Attanasio, O. P. and G. Weber (1995): "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 1121–1157.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the Macroeconomy, and Asset Prices," *Journal of Finance*, 69, 2471–2511.

Bansal, R., D. Kiku, and A. Yaron (2010): "Long-Run Risks, the Macroeconomy, and Asset Prices," *American Economic Review*, 100, 542–546.

——— (2012a): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2012b): "Risks For the Long Run: Estimation with Time Aggregation," Available at NBER: `http://www.nber.org/papers/w18305.pdf`.

Bansal, R. and I. Shaliastovich (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Barro, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *The Quarterly Journal of Economics*, 121, 823–866.

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bhamra, H. S. and R. Uppal (2014): "Asset Prices with Heterogeneity in Preferences and Beliefs," *Review of Financial Studies*, 27, 519–580.

Bidarkota, P. V. and J. McCulloch (2003): "Consumption Asset Pricing with Stable Shocks—Exploring a Solution and its Implications for Mean Equity Returns," *Journal of Economic Dynamics and Control*, 27, 399 – 421.

Blume, L. and D. Easley (2006): "If You're so Smart, why Aren't You Rich? Belief Selection in Complete and Incomplete Markets," *Econometrica*, 74, 929–966.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22, 4463–4492.

Bollerslev, T., L. Xu, and H. Zhou (2015): "Stock Return and Cash Flow Predictability: The Role of Volatility Risk," *Journal of Econometrics*, 187, 458–471.

Borovička, J. (2015): "Survival and Long-Run Dynamics with Heterogeneous Beliefs Under Recursive Preferences," Available at SSRN: `http://ssrn.com/abstract=2023501`.

Branger, N., I. Dumitrescu, V. Ivanova, and C. Schlag (2011): "Preference Heterogeneity and Survival in Long-Run Risk Models," Working Paper, Goethe University.

Burnside, C. (1998): "Solving Asset Pricing Models with Gaussian Shocks," *Journal of Economic Dynamics and Control*, 22, 329–340.

Caldara, D., J. Fernandez-Villaverde, J. Rubio-Ramirez, and W. Yao (2012): "Computing DSGE Models with Recursive Preferences and Stochastic Volatility," *Review of Economic Dynamics*, 15, 188–206.

Calin, O. L., Y. Chen, T. F. Cosimano, and A. A. Himonas (2005): "Solving Asset Pricing Models when the Price–Dividend Function Is Analytic," *Econometrica*, 73, 961–982.

Campbell, H. and J. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Campbell, J. Y. (1993): "Intertemporal Asset Pricing without Consumption Data," *The American Economic Review*, 83, 487–512.

——— (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104, 298–345.

Campbell, J. Y. and R. J. Shiller (1988a): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

——— (1988b): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

——— (1988c): "Stock Prices, Earnings, and Expected Dividends," *The Journal of Finance*, 43, 661–676.

Chen, Y., T. F. Cosimano, and A. A. Himonas (2008): "Analytic Solving of Asset Pricing Models: The by Force of Habit Case," *Journal of Economic Dynamics and Control*, 32, 3631 – 3660.

——— (2014): "On Formulating and Solving Portfolio Decision and Asset Pricing Problems," in *Handbook of the Computational Economics, Volume 3*, ed. by K. Schmedders and K. L. Judd, North-Holland, Amsterdam, 161–223.

Christiano, L. and M. Eichenbaum (1990): "Unit Roots in Real GNP: Do We Know, and Do We Care?" *Carnegie Rochester Conference Series on Public Policy*, 32, 7–61.

Cochrane, J. (1992): "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies*, 5, 243–280.

Cochrane, J. H. (1991): "A Critique of the Application of Unit Root Tests," *Journal of Economic Dynamics and Control*, 15, 275–284.

Collard, F., P. Féve, and I. Ghattassi (2006): "A Note on the Exact Solution of Asset Pricing Models with Habit Persistence," *Macroeconomic Dynamics*, 10, 273–283.

Collard, F. and M. Juillard (2001): "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control*, 25, 979–999.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, 98, 519–543.

Constantinides, G. M. and A. Ghosh (2011): "Asset Pricing Tests with Long-Run Risks in Consumption Growth," *Review of Asset Pricing Studies*, 1, 96–136.

Cowles, A. and Associates (1939): *Common-Stock Indexes*, Principia Press, Bloomington, IN.

Da, Z. and H. Yun (2011): "Electricity Consumption and Asset Prices," Available at SSRN: `http://ssrn.com/abstract=1785574`.

de Groot, O. (2013): "Computing the Risky Steady State of DSGE Models," *Economics Letters*, 120, 566–569.

——— (2015): "Solving Asset Pricing Models with Stochastic Volatility," *Journal of Economic Dynamics and Control*, 52, 308–321.

DeJong, D. (1992): "Co-Integration and Trend-Stationarity in Macroeconomic Time Series," *Journal of Econometrics*, 52, 347–370.

DeJong, D., J. Nankervis, N. Savin, and C. Whiteman (1992): "Integration Versus Trend Stationarity in Time Series," *Econometrica*, 60, 423–433.

DeJong, D. and M. Ripoll (2007): "Do Self-Control Preferences Help Explain the Puzzling Behavior of Asset Prices?" *Journal of Monetary Economics*, 54, 1035–1050.

DeJong, D. and C. Whiteman (1991a): "The Case for Trend-Stationarity is Stronger Than We Thought," *Journal of Applied Econometrics*, 6, 413–421.

——— (1991b): "Reconsidering 'Trends and Random Walks in Macroeconomic Time Series'," *Journal of Monetary Economics*, 28, 221–254.

——— (1991c): "The Temporal Stability of Dividends and Stock Prices: Evidence from the Likelihood Function," *The American Economic Review*, 81, 600–617.

Dew-Becker, I., S. W. Giglio, A. Le, and M. Rodriguez (2015): "The Price of Variance Risk," Available at SSRN: `http://ssrn.com/abstract=2471079`.

Dickey, D. and W. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

Diebold, F. and A. Senhadji (1996): "The Uncertain Unit Root in Real GNP: Comment," *The American Economic Review*, 86, 1291–1298.

Drechsler, I. and A. Yaron (2011): "What's Vol Got to Do with It," *Review of Financial Studies*, 24, 1–45.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Eraker, B. (2008): "Affine General Equilibrium Models," *Management Science*, 54, 2068–2080.

Eraker, B. and I. Shaliastovich (2008): "An Equilibrium Guide to Designing Affine Pricing Models," *Mathematical Finance*, 18, 519–543.

Fama, E. F. and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

Floden, M. (2007): "A Note on the Accuracy of Markov-Chain Approximations to Highly Persistent AR(1) Processes," *Economics Letters*, 99, 516–520.

Friedman, M. (1953): *Essays in Positive Economics*, Chicago: University of Chicago Press.

Galindev, R. and D. Lkhagvasuren (2010): "Discretization of Highly-Persistent Correlated AR(1) Shocks," *Journal of Economic Dynamics and Control*, 34, 1260–1276.

Gelfand, I. M. and S. V. Fomin (1963): *Calculus of Variations*, Prentice Hall, Englewood Cliffs.

Grossman, S. J. and R. J. Shiller (1981): "The Determinants of the Variability of Stock Market Prices," *The American Economic Review*, 71, 222–227.

Guvenen, F. (2009): "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica*, 77, 1711–1740.

Hansen, L. P., J. Heaton, J. Lee, and N. Roussanov (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, Elsevier, vol. 6, Part A, Chapter 61, 3967–4056.

Hansen, L. P., J. Heaton, and N. Li (2008): "Consumption Strikes Back?: Measuring Long Run Risk," *Journal of Political Economy*, 116, 260–302.

Hansen, L. P., J. Heaton, and A. Yaron (1996): "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics*, 14, 262–280.

Hansen, L. P. and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

Hansen, L. P. and J. A. Scheinkman (2012): "Recursive Utility in a Markov Environment with Stochastic Growth," *Proceedings of the National Academy of Sciences*, 109, 11967–11972.

Hansen, L. P. and K. J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

Heaton, J. and D. Lucas (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, 104, 443–487.

——— (2000): "Portfolio Choice in the Presence of Background Risk," *The Economic Journal*, 110, 1–26.

Judd, K. (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58, 410–452.

——— (1996): "Approximation, Perturbation, and Projection Methods in Economic Analysis," in *Handbook of the Computational Economics, Volume 1*, ed. by H. Amman, D. Kendrick, and J. Rust, North-Holland, Amsterdam, 511–585.

——— (1998): *Numerical Methods in Economics*, The MIT Press: Cambridge, Massachusetts.

Judd, K., F. Kubler, and K. Schmedders (2003): "Asset Trading Volume with Dynamically Complete Markets and Heterogeneous Agents," *The Journal of Finance*, 58, 2203–2217.

Juillard, M. (2011): "Local Approximation of DSGE Models Around the Risky Steady State," Available at IDEAS: `http://ideas.repec.org/p/ter/wpaper/0087.html`.

Kaltenbrunner, G. and L. A. Lochstoer (2010): "Long-Run Risk Through Consumption Smoothing," *Review of Financial Studies*, 23, 3190–3224.

Kendrick, J. W. (1961): *Productivity Trends in the United States*, National Bureau of Economic Research, Inc.

Kocherlakota, N. (1990): "On the 'Discount' Factor in Growth Economies," *Journal of Monetary Economics*, 25, 43–47.

Kogan, L. and R. Uppal (2001): "Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies," Available at NBER: `http://www.nber.org/papers/w8609`.

Koijen, R. S. J., H. Lustig, S. Van Nieuwerburgh, and A. Verdelhan (2010): "Long-Run Risk, the Wealth-Consumption Ratio, and the Temporal Pricing of Risk," *American Economic Review*, 100, 552–556.

Kopecky, K. and R. Suen (2010): "Finite State Markov-Chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 13, 701–714.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178.

Longstaff, F. and M. Piazzesi (2004): "Corporate Earnings and the Equity Premium," *Journal of Financial Economics*, 74, 401–421.

Lontzek, T. S., D. Narita, and O. Wilms (2015): "Stochastic Integrated Assessment of Ecosystem Tipping Risk," Working Paper, University Zurich.

Lucas, R. E. J. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Marinacci, M. and L. Montrucchio (2010): "Unique Solutions for Stochastic Recursive Utilities," *Journal of Economic Theory*, 145, 1776–1804.

Mehra, R. (2006): "The Equity Premium Puzzle: A Review," *Foundations and Trends in Finance*, 2, 1–81.

Mehra, R. and E. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Meyer-Gohde, A. (2014): "Risky Linear Approximations," Available at IDEAS: `http://ideas.repec.org/p/hum/wpaper/sfb649dp2014-034.html`.

Nakamura, E., J. Steinsson, R. Barro, and J. Ursúa (2013): "Crises and Recoveries in an Empirical Model of Consumption Disasters," *American Economic Journal: Macroeconomics*, 5, 35–74.

Nelson, C. and C. Plosser (1982): "Trends and Random Walks in Macroeconmic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–162.

Perron, P. (1989): "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 1361–1401.

Phillips, P. and P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335–346.

Piazzesi, M., M. Schneider, and S. Tuzel (2007): "Housing, Consumption and Asset Pricing," *Journal of Financial Economics*, 83, 531–569.

Pohl, W., K. Schmedders, and O. Wilms (2014): "Asset Prices with Temporary Shocks to Consumption," Available at SSRN: `http://ssrn.com/abstract=2450090`.

Qiao, X. (2013): "Cross-Sectional Evidence in Consumption Mismeasurement," Available at SSRN: `http://ssrn.com/abstract=2317486`.

Reich, G. and O. Wilms (2015): "Adaptive Grids for the Estimation of Dynamic Models," Available at SSRN: `http://ssrn.com/abstract=2650994`.

Rietz, T. A. (1988): "The Equity Risk Premium a Solution," *Journal of Monetary Economics*, 22, 117 – 131.

Rodriguez, J. (2006): "Consumption, the Persistence of Shocks, and Asset Price Volatility," *Journal of Monetary Economics*, 53, 1741–1760.

Rouwenhorst, G. (1995): *Asset Pricing Implications of Equilibrium Business Cycle Models*, Cooley, T. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ, 294–330.

Rudebusch, G. (1993): "The Uncertain Unit Root in Real GNP," *The American Economic Review*, 83, 264–272.

Sandroni, A. (2000): "Do Markets Favor Agents Able to Make Accurate Predictions?" *Econometrica*, 68, 1303–1341.

Santos, T. and P. Veronesi (2006): "Labor Income and Predictable Stock Returns," *The Review of Financial Studies*, 19, 1–44.

——— (2010): "Habit Formation, the Cross Section of Stock Returns and the Cash-Flow Risk Puzzle," *Journal of Financial Economics*, 98, 385–413.

Savov, A. (2011): "Asset Pricing with Garbage," *The Journal of Finance*, 66, 177–201.

Schorfheide, F., D. Song, and A. Yaron (2014): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Available at NBER: `http://www.nber.org/papers/w20303`.

Schwert, G. W. (2002): "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business & Economic Statistics*, 20, 5–17.

Segal, G., I. Shaliastovich, and A. Yaron (2015): "Good and Bad Uncertainty: Macroeconomic and Financial Market Implications," *Journal of Financial Economics*, 117, 369–397.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71, 421–436.

TALLARINI, T. (2000): "Risk-Sensitive Real Business Cycles," *Journal of Monetary Economics*, 45, 507–532.

TARSKI, A. (1955): "A Lattice-theoretical Fixpoint Theorem and its Applications," *Pacific Journal of Mathematics*, 5, 285–309.

TAUCHEN, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.

——— (2011): "Stochastic Volatility in General Equilibrium," *Quarterly Journal of Finance*, 1, 707–731.

TAUCHEN, G. AND R. HUSSEY (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59, 317–396.

TRIPLETT, J. E. (1997): "Measuring Consumption: The Post-1973 Slowdown and the Research Issues," *Federal Reserve Bank of St. Louis Review.*

TSIONAS, E. G. (2003): "Exact Solution of Asset Pricing Models with Arbitrary Shock Distributions," *Journal of Economic Dynamics and Control*, 27, 843 – 851.

WACHTER, J. A. (2006): "A Consumption-Based Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 79, 365–399.

WARREN, G. F. AND F. A. PEARSON (1935): *Gold and Prices*, J. Wiley and Sons.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

——— (1990): "Nonexpected Utility in Macroeconomics." *The Quarterly Journal of Economics*, 105, 29–42.

YOGO, M. (2004): "Estimating the Elasticity of Intertemporal Substitution When Instruments are Weak," *The Review of Economics and Statistics*, 86, 797–810.

ZHOU, G. AND Y. ZHU (2015): "Macroeconomic Volatilities and Long-Run Risks of Asset Prices," *Management Science*, 61, 413–430.

Curriculum Vitae of Ole Wilms

Personal Information

Citizenship	Germany Advisor Prof. Karl Schmedders
Birth date	August 16, 1987 in Eckernförde, Germany

Education

10/2012 – 04/2016	Ph.D. Studies in Management and Economics, with distinction (summa cum laude), University of Zurich Advisors: Prof. Karl Schmedders, Prof Felix Kübler
09/2014 – 12/2014	Visiting Scholar, Department of Economics, Stanford University
09/2014	Princeton Initiative: Macro, Money and Finance, Princeton University
10/2010 – 10/2012	M.Sc., Quantitative Finance, Christian-Albrechts-Universität zu Kiel
09/2011	The 2011 Zurich Initiative on Computational Economics, University of Zurich
02/2011 – 06/2011	Exchange Student, Quantitative Finance, University of Zurich
10/2007 – 10/2010	B.Sc., Economics, Christian-Albrechts-Universität zu Kiel
08/2009 – 01/2010	Exchange Student, Economics, University of California, San Diego